

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 10, 2005, 23:23:15 ; Search time 241 Seconds

(Without alignments)
1593.148 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027
Sequence: 1 gggcgaagtggtggcaaaaa.....atttaaaaaatgcaaaagct 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA.New:*
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	36.8	3.6	600	6 US-10-750-185-1996	Sequence 1996, Ap
C 2	34.6	3.4	1561	6 US-10-750-185-40906	Sequence 40906, A
C 3	34	3.3	905	6 US-10-750-185-53770	Sequence 53770, A
C 4	34	3.3	1173	6 US-10-750-185-52322	Sequence 52322, A
C 5	34	3.3	130837	7 US-11-121-086-97	Sequence 97, Appl
C 6	34	3.3	191797	7 US-11-121-086-13	Sequence 13, Appl
C 7	33.8	3.3	96988	7 US-11-117-187-196	Sequence 156, Appl
C 8	33.6	3.3	182303	7 US-11-121-086-45	Sequence 45, Appl
C 9	33.4	3.3	1515	6 US-10-750-185-40420	Sequence 40420, A
C 10	33.4	3.3	1567	6 US-10-750-185-25199	Sequence 25199, A
C 11	33.2	3.2	40439	6 US-10-993-509-1	Sequence 1, Appl
C 12	33	3.2	1128	6 US-10-750-185-54062	Sequence 54062, A
C 13	33	3.2	108214	7 US-11-117-187-211	Sequence 211, App
C 14	32.8	3.2	600	6 US-10-750-185-102	Sequence 102, App
C 15	32.8	3.2	1037	6 US-10-518-753-14	Sequence 14, Appl
C 16	32.8	3.2	1449	6 US-10-750-185-53517	Sequence 53517, A
C 17	32.8	3.2	2767	6 US-10-750-185-24844	Sequence 24844, A
C 18	32.8	3.2	4497	6 US-10-518-753-15	Sequence 15, Appl
C 19	32.8	3.2	162173	7 US-11-121-086-72	Sequence 72, Appl
C 20	32.4	3.2	161874	7 US-11-121-086-75	Sequence 75, Appl
C 21	32.2	3.1	600	6 US-10-750-185-20145	Sequence 20145, A
C 22	32.2	3.1	944	6 US-10-750-185-39148	Sequence 39148, A
C 23	32.2	3.1	2069	6 US-10-750-185-49897	Sequence 49897, A

24	32.2	3.1	186442	7 US-11-121-086-104	Sequence 104, App
C 25	32	3.1	1707	6 US-10-750-185-40279	Sequence 40279, A
C 26	31.8	3.1	4265	7 US-11-044-051-74	Sequence 74, Appl
C 27	31.8	3.1	260209	6 US-10-933-025-23	Sequence 23, Appl
C 28	31.6	3.1	319608	7 US-11-145-703-1	Sequence 1, Appl
C 29	31.4	3.1	319	7 US-11-108-172-339	Sequence 339, App
C 30	31.4	3.1	1096	6 US-10-750-185-25016	Sequence 25016, A
C 31	31.4	3.1	1441	6 US-10-750-185-48298	Sequence 48298, A
C 32	31.4	3.1	1564	6 US-10-750-185-48597	Sequence 48597, A
C 33	31.4	3.1	1838	6 US-10-750-185-51829	Sequence 51829, A
C 34	31.2	3.0	1034	6 US-10-750-185-40337	Sequence 40337, A
C 35	31.2	3.0	2744	6 US-10-750-185-30635	Sequence 30635, A
C 36	31.2	3.0	3155	6 US-10-750-185-59681	Sequence 59681, A
C 37	31.2	3.0	176760	7 US-11-121-086-51	Sequence 51, Appl
C 38	31	3.0	817	6 US-10-750-185-53425	Sequence 53425, A
C 39	31	3.0	1077	6 US-10-750-185-51726	Sequence 51726, A
C 40	31	3.0	3755	6 US-10-793-626-3642	Sequence 3642, Ap
C 41	31	3.0	4070	6 US-10-750-185-53698	Sequence 53698, A
C 42	30.8	3.0	154548	7 US-11-121-086-33	Sequence 33, Appl
C 43	30.8	3.0	1025	6 US-10-750-185-44629	Sequence 44629, A
C 44	30.8	3.0	2511	6 US-10-750-185-41003	Sequence 41003, A
C 45	30.8	3.0	3272	6 US-10-750-185-24646	Sequence 24646, A

ALIGNMENTS

RESULT 1
US-10-750-185-1996/c
; Sequence 1996, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1996
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MM118519
; US-10-750-185-1996

Query Match 3.6%; Score 36.8; DB 6; Length 600;
Best Local Similarity 52.0%; Pred. No. 0.17;
Matches 106; Conservative 0; Mismatches 97; Indels 1; Gaps 1;

163 GAGAAATTTTCCCTGATGTTAATGATCATTGGCAAGAAAGAGCAACAGATT 222
528 GATTAATTTTCAAGCTATATGATGAATAGCCCTTGTGTTAAAGTAATGTAATGTTT 469
223 CTGGATCCAGAGGGAACACATGACATTAAGCAACATACATGAAATTTAGCTGG 282
468 ATAGGTCTATTATTAAGAAAAATCTATACACACATCAACCCAGATCTTATGACA 409
283 TTA-ACGGTCAGAAAAGTCATGACAAAGAACAGATGATGTCAGACATGAGAA 341
408 TTATTCCTTGATGAATTAAGATTAACGTTTAAAAAAGAAAAAGTGAATTAAGAGAAA 349
342 TTAATTAAGGAGAGTTGATCAAGA 365
348 TATTAAGACGTAAACGTTCAAGA 325

```

RESULT 2
US-10-750-185-40906/c
: Sequence 40906, Application US/10750185
: Publication No. US20050260603A1
: GENERAL INFORMATION:
: APPLICANT: MMI GENOMICS, INC.
: APPLICANT: DENISE, Sue K.
: APPLICANT: KERR, Richard
: APPLICANT: ROSENFELD, David
: APPLICANT: HOLM, Tom
: APPLICANT: BATES, Stephen
: APPLICANT: FANTIN, Dennis
: TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
: FILE REFERENCE: MM11100-2
: CURRENT APPLICATION NUMBER: US/10/750,185
: CURRENT FILING DATE: 2003-12-31
: PRIOR APPLICATION NUMBER: US 60/437,482
: PRIOR FILING DATE: 2002-12-31
: NUMBER OF SEQ ID NOS: 64922
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 40906
: LENGTH: 1561
: TYPE: DNA
: ORGANISM: Bovine 19866881662064
US-10-750-185-40906

```

Query Match	3.4%	Score 34.6	DB 6	Length 1561
Best Local Similarity	61.8%	Pred. No. 1.5		
Matches	55	Conservative	0	Mismatches 34
			Indels 0	Gaps 0
QY	655	TTGGGCTTCTTCTTGGGGTTGGGCATTTCAGTTCATGNGTGTACTATTCATCT	714	
Db	1426	TTGGCCATGTCACATGGCATGTGGGATCTTAGTTCCTCGTGTATTATTAATCACTCAGG	1367	
QY	715	ATTGTATTAACGTTTCAACACAGTGGGC	743	
Db	1366	TGTATCTGACTCTTTGGCATCTGTGGGC	1338	

```

* RESULT 3
* US-10-750-185-53770
* Sequence: 53770, Application US/10750185
* Publication No. US20050260603A1
* GENERAL INFORMATION:
* APPLICANT: MMI GENOMICS, INC.
* APPLICANT: DENISE, Sue K.
* APPLICANT: KERR, Richard
* APPLICANT: ROSENFELD, David
* APPLICANT: HOLM, Tom
* APPLICANT: BATES, Stephen
* APPLICANT: FANTIN, Dennis
* TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
* FILE REFERENCE: MM1100-2
* CURRENT APPLICATION NUMBER: US/10/750,185
* CURRENT FILING DATE: 2003-12-31
* PRIOR APPLICATION NUMBER: US 60/437,482
* PRIOR FILING DATE: 2002-12-31
* NUMBER OF SEQ ID NOS: 64922
* SOFTWARE: PatentIn version 3.1
* SEQ ID NO 53770
* LENGTH: 905
* TYPE: DNA
* ORGANISM: Bovine 19866880693972
* US-10-750-185-53770

```

Query Match	3.3%	Score 34	DB 6	Length 905
Best Local Similarity	50.0%	Pred. No. 1.6		
Matches	85	Conservative	0	Mismatches 85; Indels 0; Gaps 0
Oy	855	ATATGTTGACATCTCTGCGGCTTCTTACGCTTCTCTCTTATGTTCTTTATCATCATGAT	914	
Db	456	ATATGCTTGAACATCTTCACAGGCTGCGAGGCTCTCTCTCATCGGGTCATATGTTAATGAAGA	515	

Qy	915	AACTGCGCTGAAGACCTTTTCATTTTACAGCCCTGAGACAGCTCTTTCGCTAGTGAATT	974
Db	516	CATTTCGCTTCCCCCTCCTACTGTTTGGCTGTGCTGATTTTCTATTTTTCCTAATAAAC	575
Qy	975	AATGCGATGCTTTTTCCCTAATAAGCAAAATTAATTTAAAAAAATGAAA	1024
Db	576	ATCTTAATTTTTTTTATATAGCAAAAGAAAAATTTTAAATTAAAAAAA	625

```

1 RESULT 4
2 US-10-750-185-52322/c
3 Sequence 52322, Application US/10750185
4 Publication NO. US20050260603A1
5 GENERAL INFORMATION:
6 APPLICANT: MMI GENOMICS, INC.
7 APPLICANT: DENISE, Sue K.
8 APPLICANT: KERR, Richard
9 APPLICANT: ROSENFELD, David
10 APPLICANT: HOLM, Tom
11 APPLICANT: BATES, Stephen
12 APPLICANT: FANTIN, Dennis
13 TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
14 FILE REFERENCE: MM1100-2
15 CURRENT APPLICATION NUMBER: US/10/750,185
16 CURRENT FILING DATE: 2003-12-31
17 PRIOR APPLICATION NUMBER: US 60/437,482
18 PRIOR FILING DATE: 2002-12-31
19 NUMBER OF SEQ ID NOS: 64922
20 SOFTWARE: PatentIn version 3.1
21 SEQ ID NO 52322
22 LENGTH: 1173
23 TYPE: DNA
24 ORGANISM: Bovine 19866880907657
25 US-10-750-185-52322

```

Query Match	3.3%	Score 34;	DB 6;	Length 1173;
Best Local Similarity	66.2%	Pred. No. 1.9;		
Matches 49;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;

Qy	953	AGCTCTCTTCCTGAGTGAATTAATGCGTGTCTTTCGGAATPACGAAATAAATTTA	1012
Db	231	ATTAGTATTACTGAGTGAATTAATGTAAGTGTGTCCTATACACATACATTTA	172
Qy	1013	AAAAAAAAAAGT 1026	
Db	171	CCAAAAAATAGT 158	

```

RESULT 5
US-11-121-086-97/c
; Sequence 97, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: NIELSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121.086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ. ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ. ID NO. 97
; LENGTH: 170837
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-97

```

Query Match	3.3%	Score 34	DB 7	Length 170837
Best Local Similarity	52.1%	Pred. No. 40		
Matches 76	Conservative 0	Mismatches 70	Indels 0	Gaps 0

RESULT 9
US-10-750-185-40420
; Sequence 40420, Application US/10750185
; Publication No. US20050260603A1

```

; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40420
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Bovine 19866881526906
; US-10-750-185-40420

```

```

Query Match      3.3% Score 33.4; DB 6; Length 1515;
Best Local Similarity 46.7%; Pred.No.3.4;
Matches 106; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

```

```

Qy 25 AAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAATTGATGAGATGTT 84
Db 23 AATCAGATTTCAGCAAAATAGATATTAATTAGTCAGACAGGACAAACCTTTT 82
Qy 85 TCCCCCAAGCCACTATTTTCTCTTCAATTGCTGAACAAGCTCCAGAGGCTGGA 144
Db 83 TCCCAACAACACTGGAACCTGCTTCTGCTTTAGGAACTCTGACAAGGACCTT 142
Qy 145 ACATACCTTGTCTCTGAGAAATTTTCCCTGATGTTATTAAGATACATTGGCAGAA 204
Db 143 ACATGAATCATTCATTTTGTAAATACAAATTTCTGACAGAGGCTTAGCTAGCCAGT 202
Qy 205 AAGAAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCATGAA 251
Db 203 TCGATTGAGAAACAGAAATCAGCATTAATGATGATGACACAGATGAA 249

```

```

RESULT 10
US-10-750-185-25199/c
; Sequence 25199, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25199
; LENGTH: 1567
; TYPE: DNA
; ORGANISM: Bovine 19866880415063
; US-10-750-185-25199

```

```

Query Match      3.3% Score 33.4; DB 6; Length 1567;
Best Local Similarity 47.8%; Pred.No.3.4;
Matches 97; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```

```

Qy 283 TTAAAGTCCCAAGAAAAGTCACTGACAAAGAACACAGATGTATCTTCAGACATGAGAAAT 342
Db 582 TTATGCTTGAAGAAATATATTTATTTAGAGCACAAACATGATGATGATGTAAT 523
Qy 343 AATAAAAACGAGTGTGATCAAGAAATATATCTTCTCCATTAAGAGAGATGATGACA 402
Db 522 TATCTCAATGAGAGATTTCTCAGAAAGATAGTACACACATTAATGACTATGTCACATG 463
Qy 403 ATGATCCCAAGACAAATGTTCAAAAGATGCAATGATACACTACTGCTGACAGCTACA 462
Db 462 CAATGACACAGAGATCATCTCGAAGAACATGACAGATATTCAGAGACGATCATGAAAAAA 403
Qy 463 AACACCTCTGCATATTACATGTA 485
Db 402 AATTTTCTATTTAGCTTGAA 380

```

```

RESULT 11
US-10-993-509-1
; Sequence 1, Application US/10993509
; Publication No. US20050250121A1
; GENERAL INFORMATION:
; APPLICANT: Aerresens, Jercoen
; APPLICANT: Athanasiou, Maria
; APPLICANT: Brain, Carlos
; APPLICANT: Cohen, Nadine
; APPLICANT: Dain, Bradley
; APPLICANT: Denton, R. Rex
; APPLICANT: Judson, Richard S.
; APPLICANT: Ozdemir, Vural
; APPLICANT: Reed, Carol R.
; TITLE OF INVENTION: NTRK Genetic Markers Associated with Progression of Alzheimer's
; FILE REFERENCE: 2300.0060001
; CURRENT APPLICATION NUMBER: US/10/993,509
; PRIOR FILING DATE: 2004-11-22
; PRIOR APPLICATION NUMBER: US 60/524,637
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 40439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2402)..(2402)
; OTHER INFORMATION: n is 'c' or 'g'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2722)..(2722)
; OTHER INFORMATION: n is 'c' or 't'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2799)..(2799)
; OTHER INFORMATION: n is 'g' or 'a'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3795)..(3898)
; OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6337)..(6436)
; OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11729)..(11828)
; OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14905)..(15004)
; OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (117523)..(117622)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11942)..(11941)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (22316)..(22415)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (24686)..(24785)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (27030)..(27129)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29535)..(29634)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31929)..(32028)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (34403)..(34502)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (35929)..(37028)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; US-10-993-509-1

```

Query Match 3.2%; Score 33.2; DB 6; Length 40439;

Best Local Similarity 34.5%; Pred. No. 29; Matches 116; Conservative 0; Mismatches 218; Indels 2; Gaps 1;

```

Qy 485 ACCTCTCTCTCTCTCTCAAGAGTGTCTATTTTGGCATCATCACTGCTGTCTGCTTA 544
Db 34298 ACTGCAACCTCGGCTCTCCAGGTTCAAGTATCTCCACCTCACTCCACAGTACTG 34357
Qy 545 GAAGAACGGCTTCTCTGTCATGAGAGAAATCATPAACAGCGTGGCAAGAGGCC 604
Db 34358 GATTAACAGGATGTGCAACCATGTCTGTAAAT--TTTCATATTTTNNNNNNNNNN 34415
Qy 605 ATCTTTTCCATCGGTATGTGCTCCTAGAGCGTCTTCGAGATCTAGTTGGCTTTC 664
Db 34416 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 34475
Qy 665 TTTCTGCGTTGGGCACTTTCAGTCTCATGTGTACTATTTCTATCATTTATGTAAC 724
Db 34476 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 34535
Qy 725 GGTTCCTCAACCAAGTGGGCAACAAGAACCTCACTCTGTAATPAACATGAGAAATAGCC 764
Db 34536 TGTTCCTCAACCAAGTGGGCAACAAGAACCTCACTCTGTAATPAACATGAGAAATAGCC 34595
Qy 785 ACGGCGATCTCCAGCAACCAATCTCTCATGTTTTC 820
Db 34596 ACGGCGATCTCCAGCAACCAATCTCTCATGTTTTC 34631

```

RESULT 12

```

US-10-750-185-54062
; Sequence 54062, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

```

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54062
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Bovine
; US-10-750-185-54062

```

Query Match 3.2%; Score 33; DB 6; Length 1128;

Best Local Similarity 52.6%; Pred. No. 37; Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```

Qy 274 TTTAGCTGTAAACGCTGCGAGAAAGTCACTGACAAAGAACAGATGTATGTCAGA 333
Db 708 TTTAACTGATGAGGTGGGAAATTAATCATCTAGTATTCACAAATTTAAACATTTGA 767
Qy 334 CATGAAATTAATTAACGAGATTTGATCAAGAAATATCTTCTCTCAATTAAGAGGAT 393
Db 768 CATATTTAAAGATGATGATCTGTTTAAATGTTGCTCAAAATTTACAAAG 827
Qy 394 GTCATCAATGATGCC 410
Db 828 CAGAGCAATTAATATTC 844

```

RESULT 13

US-11-117-187-211

```

; Sequence 211, Application US/11117187
; Publication No. US2005026560A1
; GENERAL INFORMATION:
; APPLICANT: PREUS, DAPHNE
; APPLICANT: COPEHAWER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; PRIOR FILING DATE: 2005-04-28
; PRIOR FILING DATE: 2000-03-17
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-11-117-187-211

```

```

Qy 159 TCTTGAAATTTTCCCTGATGTTTAAAGATACCTTGGCAAGAAAGAGCAACAC 218
Db 493328 TATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 493387
Qy 219 GATTCGATCCAGAGGAGGAAACCATGAAGATTAACGACATFATGAATTTG 278
Db 493388 AATFATCTGAATTAATCTTGAATTCATTTGTTTAACTTATTTGTTATTAATTAATAC 493447
Qy 279 CTGTTAACGTCGCAAGAAAGTCACTGCAAGAAACAGATGTATGTCAGACATGA 338

```

Query Match 3.2%; Score 33; DB 7; Length 1082144;

Best Local Similarity 43.5%; Pred. No. 1,840; Matches 150; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

Db 493448 AAGTTTCAAACTTAAACCAATTTTAGTAAAAAGTAAACGATTTATTTATTAATAATTAAA 493507

Qy 339 GAATATATAAAAACGAGTTGATCAAGAAATTAATCTTTCTCCAAATTAAGACGATGTCAAT 398

Db 493508 CAATTTTATAGTCTTTTCTTTTCTTCCAAACCTGTTCCGAAACCGATTCGAATCCATATCAT 493567

Qy 399 CACATGTGATCCCAAGACAATTTGTTCAAAAGATGAATATGATACACTCTGCTGCAGCT 458

Db 493568 CCAAAACCGAAACCAACCTGTATAGTCATATTAATCCGAACGATATTAATTAACCTTTCTATCC 493627

Qy 459 CACAACAACCTCTGCATATTATCACTGACCTCCTGCTCTCAAA 503

Db 493628 AATTTATGCAATTCAAAAAAATTGCATCTGATTCGAAATCCGAAA 493672

```

RESULT 14
US-10-750-185-102/c
; Sequence 102, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine
MMBT02133
US-10-750-185-102

```

[illegible]

```

RESULT 15
US-10-518-753-14
; Sequence 14, Application US/10518753
; Publication No. US20050262589A1
; GENERAL INFORMATION:
; APPLICANT: Calgene LLC
; TITLE OF INVENTION: Nucleic Acid Sequences and Methods of Use for the Production of Fatty Acids
; TITLE OF INVENTION: with Modified Polyunsaturated Fatty Acids
; FILE REFERENCE: 16518.129
; CURRENT APPLICATION NUMBER: US/10/518,753
; CURRENT FILING DATE: 2004-12-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Glycine max
US-10-518-753-14

```

	Query Match	3.2%	Score 32.8;	DB 6;	Length 1037;	
	Query Similarity	59.8%	Pred. No. 4.1;			
	Best Local					
	Matches	55;	Conservative	0;	Mismatches	37;
					Indels	0;
					Gaps	0;
Qy		136	AAGCGTGACATACCTTTGCTTCTTGAGAATTTTTCCGTGAATGATTATTAAGATACAT	195		
Db		39	AATGTTAGATACTACTCTTCCTCTTTTTTTTTTTTTTGGAGTTAATGCATGATA	98		
Qy		196	TGGCAGAAAAAGAAAGACACAGATTCTGGG	227		
Db		99	TAGCTAGAAAAGAAAAGAAACAATTAAG	130		

Search completed: December 11, 2005, 01:35:02
Job time : 245 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 22:15:03 ; Search time 890 Seconds

(without alignments)
9542.305 Million cell updates/sec

Title: US-10-031-158b-13

Perfect score: 1027
Sequence: 1 GGGCAAGAGTGGGCAAAA.....atttaaaaaatgaagaagtc 1027

Scoring table: IDENTITY_NUC
Gap 10.0 , Gape 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027	100.0	1027	5	US-10-205-823-404
2	1027	100.0	1027	5	US-10-756-149-2859
3	1027	100.0	1027	10	US-11-051-454-404
4	1013.8	98.7	1395	8	US-10-643-795A-23
5	1013.8	98.7	1395	6	US-10-948-518-23
6	968.2	94.3	1799	6	US-10-510-510-92
7	945.6	92.1	1418	3	US-09-957-708-16
8	906	88.2	1155	3	US-09-925-300-58
9	902.2	87.8	1421	3	US-09-954-456-317
10	902.2	87.8	1421	3	US-10-843-641A-3344
11	838	81.6	1586	3	US-09-960-706-676
12	838	81.6	1586	3	US-09-873-319-424
13	838	81.6	1586	3	US-10-847-918-14
14	795.8	77.5	1162	8	US-10-357-930-21517
15	795.8	77.5	1162	8	US-10-357-930-24142
16	795.8	77.5	1162	8	US-10-357-930-24529
17	795.8	77.5	1162	8	US-10-357-930-24533
18	795.8	77.5	1162	8	US-10-357-930-24840
19	795.8	77.5	1162	8	US-10-357-930-25151
20	795.8	77.5	1162	8	US-10-357-930-27359
21	795.8	77.5	1162	8	US-10-357-930-27740
22	619	60.3	924	8	US-10-357-930-25081
23	589	57.4	124933	9	US-10-737-082-97

24	589	57.4	124933	9	US-10-765-790-97	Sequence 97, Appl
C 25	498.8	48.6	1316	8	US-10-357-930-25082	Sequence 25082, A
C 26	498.8	48.6	1316	8	US-10-357-930-26474	Sequence 26474, A
C 27	473.2	46.1	486	8	US-10-357-930-25002	Sequence 25002, A
C 28	449.8	43.7	825	6	US-10-340-536-29	Sequence 29, Appl
C 29	446.8	43.5	784	8	US-10-357-930-21887	Sequence 21887, A
C 30	444	43.2	511	3	US-09-780-669-72	Sequence 72, Appl
C 31	444	43.2	511	3	US-09-780-669-72	Sequence 72, Appl
C 32	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 33	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 34	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 35	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 36	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 37	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 38	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 39	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 40	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 41	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 42	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 43	416.2	40.5	569	8	US-10-357-930-43494	Sequence 43494, A
C 44	410.4	40.0	539	8	US-10-357-930-22784	Sequence 22784, A
C 45	410.4	40.0	539	8	US-10-357-930-28636	Sequence 28636, A

ALIGNMENTS

```
RESULT 1
US-10-205-823-404
; Sequence 404, Application US/10205823
; Publication No. US20030108963A1
GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarepu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kametkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-404
Query Match 100.0%; Score 1027; DB 5; Length 1027;
Best Local Similarity 100.0%; Pred. No. 1.1e+293;
Matches 1027; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
1 GGGCAAGTGGGCAAAAATCAAGTATTTGTCGCGAACAAAGTTATCATTTACA 60
|||||
```

Db 1 GGGCAAGATTGGGCAAAAAATCAAGTATTGGTCCCGGACAAAGCTTATCATTAACA 60
Qy 61 GATTAACAACCTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCT 120
Db 61 GATTAACAACCTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCT 120
Qy 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGAT 180
Db 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGAT 180
Qy 181 GTTATTAAAGATACATTTGGCAAGAAAGAAAGCAACAGATTTTGGATCCCAAGAGGG 240
Db 181 GTTATTAAAGATACATTTGGCAAGAAAGAAAGCAACAGATTTTGGATCCCAAGAGGG 240
Qy 241 AACACATGAAGACTTAACGACATACATGAAATTTAGCTGGTTAACGGTCCAGAAAAG 300
Db 241 AACACATGAAGACTTAACGACATACATGAAATTTAGCTGGTTAACGGTCCAGAAAAG 300
Qy 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAACGAGATTGAT 360
Db 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAACGAGATTGAT 360
Qy 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAATGATGATCCCAAGACAT 420
Db 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAATGATGATCCCAAGACAT 420
Qy 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTAC 480
Db 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTAC 480
Qy 481 ATGTACCTCTCTGCTGCTCTCAAGATGTGTCTATTTTGGCATCATACCTGCTGTCTG 540
Db 481 ATGTACCTCTCTGCTGCTCTCAAGATGTGTCTATTTTGGCATCATACCTGCTGTCTG 540
Qy 541 CTTCGAAAGAAAGGGCTTTTCTGCTGCAATGGAGAGAAATCATTAACAGCGTGGCAAGAGA 600
Db 541 CTTCGAAAGAAAGGGCTTTTCTGCTGCAATGGAGAGAAATCATTAACAGCGTGGCAAGAGA 600
Qy 601 GGCACATCTTCTCTCATCGGTTATTTGCTTGAAGAGCTCTTCTGAAGATCTAGTTGGGC 660
Db 601 GGCACATCTTCTCTCATCGGTTATTTGCTTGAAGAGCTCTTCTGAAGATCTAGTTGGGC 660
Qy 661 TTTCTTCTGGGTTTGGGCAATTTCAATGTTCTCATGTGTGTAATACTTATCATTAATTGTA 720
Db 661 TTTCTTCTGGGTTTGGGCAATTTCAATGTTCTCATGTGTGTAATACTTATCATTAATTGTA 720
Qy 721 TAAAGGTTTCAAAACCAATGGGCAACAGAGAAACCTCACTCTGTAAATAACAATAGAGAAAT 780
Db 721 TAAAGGTTTCAAAACCAATGGGCAACAGAGAAACCTCACTCTGTAAATAACAATAGAGAAAT 780
Qy 781 AGCCACGGGCGATCTCCAGACCAATCTCTCCATGTCTTCCACAGGTCCTCCAGCCCAACC 840
Db 781 AGCCACGGGCGATCTCCAGACCAATCTCTCCATGTCTTCCACAGGTCCTCCAGCCCAACC 840
Qy 841 AAATAGCGCTGCTATGATGTAAGATCATCTGCGGCTTCTAGCTTGTCTCTTCTTAGTG 900
Db 841 AAATAGCGCTGCTATGATGTAAGATCATCTGCGGCTTCTAGCTTGTCTCTCTTCTTAGTG 900
Qy 901 TTTCTTAAATCAAGATACTGCTCGGAAAGCTTCTCATTTTAAACAGCCCTGAAGAGCTTCT 960
Db 901 TTTCTTAAATCAAGATACTGCTCGGAAAGCTTCTCATTTTAAACAGCCCTGAAGAGCTTCT 960
Qy 961 TTGCTAGTTGAATTAATGTGTGTGTTTCTCGTAAATAAGCAAAATTAATTTAAAAAATG 1020
Db 961 TTGCTAGTTGAATTAATGTGTGTGTTTCTCGTAAATAAGCAAAATTAATTTAAAAAATG 1020
Qy 1021 AAAAGTT 1027
Db 1021 AAAAGTT 1027

RESULT 2
US-10-756-149-2859

; Sequence 2859, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Azile, Nataasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2859
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-2859

Query Match 100.0%; Score 1027; DB 9; Length 1027;
Best Local Similarity 100.0%; Pred. No. 1.1e-293;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCAAGATTGGGCAAAAAATCAAGTATTGGTCCCGGACAAAGCTTATCATTAACA 60
Db 1 GGGCAAGATTGGGCAAAAAATCAAGTATTGGTCCCGGACAAAGCTTATCATTAACA 60
Qy 61 GATTAACAACCTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCT 120
Db 61 GATTAACAACCTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGCT 120
Qy 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGAT 180
Db 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGAT 180
Qy 181 GTTATTAAAGATACATTTGGCAAGAAAGAAAGCAACAGATTTTGGATCCCAAGAGGG 240
Db 181 GTTATTAAAGATACATTTGGCAAGAAAGAAAGCAACAGATTTTGGATCCCAAGAGGG 240
Qy 241 AACACATGAAGACTTAACGACATACATGAAATTTAGCTGGTTAACGGTCCAGAAAAG 300
Db 241 AACACATGAAGACTTAACGACATACATGAAATTTAGCTGGTTAACGGTCCAGAAAAG 300
Qy 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAACGAGATTGAT 360
Db 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAACGAGATTGAT 360
Qy 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAATGATGATCCCAAGACAT 420
Db 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAATGATGATCCCAAGACAT 420
Qy 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTAC 480
Db 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTAC 480
Qy 481 ATGTACCTCTCTGCTGCTCTCAAGATGTGTCTATTTTGGCATCATACCTGCTGTCTG 540
Db 481 ATGTACCTCTCTGCTGCTCTCAAGATGTGTCTATTTTGGCATCATACCTGCTGTCTG 540
Qy 541 CTTCGAAAGAAAGGGCTTTTCTGCTGCAATGGAGAGAAATCATTAACAGCGTGGCAAGAGA 600
Db 541 CTTCGAAAGAAAGGGCTTTTCTGCTGCAATGGAGAGAAATCATTAACAGCGTGGCAAGAGA 600
Qy 601 GGCACATCTTCTCTCATCGGTTATTTGCTTGAAGAGCTCTTCTGAAGATCTAGTTGGGC 660
Db 601 GGCACATCTTCTCTCATCGGTTATTTGCTTGAAGAGCTCTTCTGAAGATCTAGTTGGGC 660
Qy 661 TTTCTTCTGGGTTTGGGCAATTTCAATGTTCTCATGTGTGTAATACTTATCATTAATTGTA 720
Db 661 TTTCTTCTGGGTTTGGGCAATTTCAATGTTCTCATGTGTGTAATACTTATCATTAATTGTA 720
Qy 721 TAAAGGTTTCAAAACCAATGGGCAACAGAGAAACCTCACTCTGTAAATAACAATAGAGAAAT 780
Db 721 TAAAGGTTTCAAAACCAATGGGCAACAGAGAAACCTCACTCTGTAAATAACAATAGAGAAAT 780

QY 781 AGCGAGGATCTCCAGCAGCAATCTCCATGTTTTCCAGAGCTCCAGCAACC 840
|
|
|
Db 781 AGCGAGGATCTCCAGCAGCAATCTCCATGTTTTCCAGAGCTCCAGCAACC 840
|
|
|
QY 841 AATAGCGCTGTATAGTATAGATCTCGCGCTTCTAGCTTCTCTTATAGT 900
|
|
|
Db 841 AATAGCGCTGTATAGTATAGATCTCGCGCTTCTAGCTTCTCTTATAGT 900
|
|
|
QY 901 TTCTTATATAGATTAATGCTGAGAGCTTTTATTTTACAGCCCTGAAAGCTTTCT 960
|
|
|
Db 901 TTCTTATATAGATTAATGCTGAGAGCTTTTATTTTACAGCCCTGAAAGCTTTCT 960
|
|
|
QY 961 TTCTGTTGAATTAATGCTGAGTGTGTTTTCGTAATAGCAAAATTAATAAAATG 1020
|
|
|
Db 961 TTCTGTTGAATTAATGCTGAGTGTGTTTTCGTAATAGCAAAATTAATAAAATG 1020
|
|
|
QY 1021 AAAAGTT 1027
|
|
|
Db 1021 AAAAGTT 1027
|
|
|

RESULT 3
US-11-051-454-404
; Sequence 404, Application US/11051454
; Publication No. US20050191673A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kametkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Giacit, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/11/051,454
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-051-454-404

Query Match 100.0%; Score 1027; DB 10; Length 1027;
Best Local Similarity 100.0%; Pred. No. 1.1e-293;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAAGAGTTGGGCAAAAATCAAGTATTTGGTCCCGAGCAAAAGCTTATATTACA 60
|
|
|
Db 1 GGGCAAGAGTTGGGCAAAAATCAAGTATTTGGTCCCGAGCAAAAGCTTATATTACA 60
|
|
|
QY 61 GATAAACAACCTTGATGAGATGTTTCCCGCAAGCCACTATTTTCTTCTCAATTGCT 120
|
|
|

Db 61 GATAAACAACCTTGATGAGATGTTTCCCGCAAGCCACTATTTTCTTCTCAATTGCT 120
|
|
|
QY 121 GAAACAACCTCCAGAAAGCTGGAACATATCTTTGCTCTTGAAGAAATTTTCCCTGAT 180
|
|
|
Db 121 GAAACAACCTCCAGAAAGCTGGAACATATCTTTGCTCTTGAAGAAATTTTCCCTGAT 180
|
|
|
QY 181 GTTATTATAGATCATATGGAAGAAAGAGAGCAACAGCTTCTGGATCCAGAGAGG 240
|
|
|
Db 181 GTTATTATAGATCATATGGAAGAAAGAGAGCAACAGATTTCTGGATCCAGAGAGG 240
|
|
|
QY 241 AACACCATGAAAGCTATACGACATATCATGAAATTTAGTGTGTTAAAGGTCCAGAAAG 300
|
|
|
Db 241 AACACCATGAAAGCTATACGACATATCATGAAATTTAGTGTGTTAAAGGTCCAGAAAG 300
|
|
|
QY 301 TCACTGAGCAAGAAACACAGATGTATGCTCAGACATGAGATTAATAAAACGAGTTGAT 360
|
|
|
Db 301 TCACTGAGCAAGAAACACAGATGTATGCTCAGACATGAGATTAATAAAACGAGTTGAT 360
|
|
|
QY 361 CAAGAAATTTATCTTCTCCTCAATTAAGAGAGATGTATCATCAATGATCCCAAGCAAT 420
|
|
|
Db 361 CAAGAAATTTATCTTCTCCTCAATTAAGAGAGATGTATCATCAATGATCCCAAGCAAT 420
|
|
|
QY 421 TGTTCAAAAGATGCAAAATGATACACTATGCTGAGCTGACAAACAACCTTGCATATTAC 480
|
|
|
Db 421 TGTTCAAAAGATGCAAAATGATACACTATGCTGAGCTGACAAACAACCTTGCATATTAC 480
|
|
|
QY 481 ATGTACCTCTCTCTGCTCTCTCAAGAGTGTGCTATTTTGGCATCATCTGCTGCTG 540
|
|
|
Db 481 ATGTACCTCTCTCTGCTCTCTCAAGAGTGTGCTATTTTGGCATCATCTGCTGCTG 540
|
|
|
QY 541 CTTAGAAAGAACGGCTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGCAAGA 600
|
|
|
Db 541 CTTAGAAAGAACGGCTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGCAAGA 600
|
|
|
QY 601 GGCCATCTTTTCTCTCATCGGTTATTTGCTCTAGAACGCTTCTGAGAGCTTATGTTGGC 660
|
|
|
Db 601 GGCCATCTTTTCTCTCATCGGTTATTTGCTCTAGAACGCTTCTGAGAGCTTATGTTGGC 660
|
|
|
QY 661 TTTCTTTCTGGGTTTGGGCAATTTCACTTCTCATGCTGATCTATTTATTTATTTGTA 720
|
|
|
Db 661 TTTCTTTCTGGGTTTGGGCAATTTCACTTCTCATGCTGATCTATTTATTTATTTGTA 720
|
|
|
QY 721 TAACGTTTTCAAACAGATGGGCAACAGAGAACTCATCTGTAATTAACATGAGGAAT 780
|
|
|
Db 721 TAACGTTTTCAAACAGATGGGCAACAGAGAACTCATCTGTAATTAACATGAGGAAT 780
|
|
|
QY 781 AGCGAGGATCTCCAGCAGCAATCTCTCATGTTTTCCAGAGCTCTCCAGCAACC 840
|
|
|
Db 781 AGCGAGGATCTCCAGCAGCAATCTCTCATGTTTTCCAGAGCTCTCCAGCAACC 840
|
|
|
QY 841 AATAGCGCTGTATAGTATAGATCTCGCGCTTCTAGCTTCTCTTATAGT 900
|
|
|
Db 841 AATAGCGCTGTATAGTATAGATCTCGCGCTTCTAGCTTCTCTTATAGT 900
|
|
|
QY 901 TTCTTATATAGATTAATGCTGAGAGCTTTTATTTTACAGCCCTGAAAGCTTTCT 960
|
|
|
Db 901 TTCTTATATAGATTAATGCTGAGAGCTTTTATTTTACAGCCCTGAAAGCTTTCT 960
|
|
|
QY 961 TTCTGTTGAATTAATGCTGAGTGTGTTTTCGTAATAGCAAAATTAATAAAATG 1020
|
|
|
Db 961 TTCTGTTGAATTAATGCTGAGTGTGTTTTCGTAATAGCAAAATTAATAAAATG 1020
|
|
|
QY 1021 AAAAGTT 1027
|
|
|
Db 1021 AAAAGTT 1027
|
|
|

RESULT 4
US-10-643-795A-23
; Sequence 23, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:

Query Match 98.7%; Score 1013.8; DB 9; Length 1395;
Best Local Similarity 99.8%; Pred. No. 1.1e-289;
Matches 1015; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 4 CAAGAGTTGGGCAAAAAATCAAGATATTTGGTCCCGAAACAAAGCTTATCATACAT 63
DB 379 CAAGAGTTGGGCAAAAAATCAAGATATTTGGTCCCGAAACAAAGCTTATCATACAT 438
QY 64 AAACAATTGATGAGATGTTTCCCGAAGCCCACTATTTTCTTCTTCAATTGGCGAA 123
DB 439 AAACAATTGATGAGATGTTTCCCGAAGCCCACTATTTTCTTCTTCAATTGGCGAA 498
QY 124 ACAAGCTCCAGAGGCTGGAACAATACCTTGTCTTCTGAGAAATTTTCCCTGATGT 183
DB 499 ACAAGCTCCAGAGGCTGGAACAATACCTTGTCTTCTGAGAAATTTTCCCTGATGT 558
QY 184 ATTAAGATACATTTGGCAAGAAAGAGCAACGATTTGGGATCCAGAGGGGAAC 243
DB 559 ATTAAGATACATTTGGCAAGAAAGAGCAACGATTTGGGATCCAGAGGGGAAC 618
QY 244 ACCATGAAGCTAACGACATACATGAATTTAGCTGTTAACGGTGGCAGAAAGTCA 303
DB 619 ACCATGAAGCTAACGACATACATGAATTTAGCTGTTAACGGTGGCAGAAAGTCA 678
QY 304 CTGGACAAGAACACAGATGATCGTCAGACATGAGATTAATAAAACGAGTTGATCAA 363
DB 679 CTGGACAAGAACACAGATGATCGTCAGACATGAGATTAATAAAACGAGTTGATCAA 738
QY 364 GAAATTAATCTTCTCCATTAAGAGCGATGTCATCAATGATCCCAAGACATTTGT 423
DB 739 GAAATTAATCTTCTCCATTAAGAGCGATGTCATCAATGATCCCAAGACATTTGT 798
QY 424 TCAAAAATGCAATGATACCTACTGCTGAGAGTCAAAAACCTTGCATATTAATG 483
DB 799 TCAAAAATGCAATGATACCTACTGCTGAGAGTCAAAAACCTTGCATATTAATG 858
QY 484 TACCTCTCTGCTCTCAAGAGTGTGATTTTGCATCATCCTGCTGCTCT 543
DB 859 TACCTCTCTGCTCTCAAGAGTGTGATTTTGCATCATCCTGCTGCTCT 918
QY 544 AGAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGGCAAGAGGC 603
DB 919 AGAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGGCAAGAGGC 978
QY 604 CATCTTTTCTCATCGGTATATTCCTTAAGAAAGCTCTTCTGAGGATCTAGTTGGCTTT 663
DB 979 CATCTTTTCTCATCGGTATATTCCTTAAGAAAGCTCTTCTGAGGATCTAGTTGGCTTT 1038
QY 664 CTTTCTGGGTTGGGCAATTTCACTTCAATGATGATATTTATCATATTTATTTAA 723
DB 1039 CTTTCTGGGTTGGGCAATTTCACTTCAATGATGATATTTATCATATTTATTTAA 1098
QY 724 CGGTTTTCAACCAAGTGGGCAACAGAGAACTCACTCTGTAATTAACAATGAGAAATAGC 783
DB 1099 CGGTTTTCAACCAAGTGGGCAACAGAGAACTCACTCTGTAATTAACAATGAGAAATAGC 1158
QY 784 CACGGCATCTCCAGACCAATCTCTCATATTTTCCACAGCTCTTCAAGCAACCCAAA 843
DB 1159 CACGGCATCTCCAGACCAATCTCTCATATTTTCCACAGCTCTTCAAGCAACCCAAA 1218
QY 844 TAGGCGCTGATATGATGATGATCTGCGGCTCTAGACCTTGTCCCTCTTAGTGTTC 903
DB 1219 TAGGCGCTGATATGATGATGATCTGCGGCTCTAGACCTTGTCCCTCTTAGTGTTC 1278
QY 904 TTTAATCAGATTAATGCTGGAAACCTTTCAATTTAACAAGCCCTGAAGACATCTTTT 963
DB 1279 TTTAATCAGATTAATGCTGGAAACCTTTCAATTTAACAAGCCCTGAAGACATCTTTT 1338
QY 964 CTAGTTGAATTAATGATGATGATTTTCCGTAATAAGCAAAATTAATTTAAAAAATG 1020
DB 1339 CTAGTTGAATTAATGATGATGATTTTCCGTAATAAGCAAAATTAATTTAAAAAATG 1395
```

RESULT 6
US-10-101-510-492
; Sequence 492; Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT FILING DATE: 2002-03-20
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 492
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-492

Query Match 94.3%; Score 968.2; DB 6; Length 1799;
Best Local Similarity 98.1%; Pred. No. 4.3e-276;
Matches 990; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

```
QY 20 AAATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATTAACAATTGATGACAG 79
DB 791 ATAAGAACTCTTTGGCAGTGAACAACACTGTTGTACAGATTAACAACCTGATGACAG 850
QY 80 ATGTTTCCCGCAAGCCCACTATTTTCTTCTTCAATGCTGAACAAAGCTCCAGAGG 139
DB 851 ATGTTTCCCGCAAGCCCACTATTTTCTTCTTCAATGCTGAACAAAGCTCCAGAGG 910
QY 140 CTGGAACATACCTTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTAATTAATGATGCGC 199
DB 911 CTGGAACATACCTTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTAATTAATGATGCGC 970
QY 200 AAGAAAAGAGCAACAGATTTCT-GGGATCCAGAGGGGAAACCATGAAGACTAAC 258
DB 971 AAGAAAAGAGCAACAGATTTCTGGGATCCAGAGGGGAAACCATGAAGACTAAC 1030
QY 259 GACACATACATGAATTTAGCTGTTAACGGTCCAGAAAAGTCACTGACAAAGAAC 318
DB 1031 GACACATACATGAATTTAGCTGTTAACGGTCCAGAAAAGTCACTGACAAAGAAC 1090
QY 319 AGATGATCTGTCAGACATGAATTAATAAAGAGTGTATGAAGAAATTAATCTTCT 378
DB 1091 AGATGATCTGTCAGACATGAATTAATAAAGAGTGTATGAAGAAATTAATCTTCT 1150
QY 379 CCAATTAAGCGGATGTCATCAATGATCCCAAGAACAAATGTTCAAAAGATGCAAT 438
DB 1151 CCAATTAAGCGGATGTCATCAATGATCCCAAGAACAAATGTTCAAAAGATGCAAT 1210
QY 439 GATACACTACTGCTGAGCTCACAAACACTCTGCAATTAATGATGATCCTCTGCTC 498
DB 1211 GATACACTACTGCTGAGCTCACAAACACTCTGCAATTAATGATGATCCTCTGCTC 1270
QY 499 CTGAAGATGATGATTTTGTGCAATCATCCTGCTGTCTGTTAGAAAGAGGCTTTC 558
DB 1271 CTGAAGATGATGATTTTGTGCAATCATCCTGCTGTCTGTTAGAAAGAGGCTTTC 1330
QY 559 TGTGCAATGGAAGAAATCATTAACAGAGGTTGGCAAGAGAGCCATCTTTCTCATC 618
DB 1331 TGTGCAATGGAAGAAATCATTAACAGAGGTTGGCAAGAGAGCCATCTTTCTCATC 1390
QY 619 GGTATATGTCCTAGAGGCTCTTCTGAGGATCTAGTTGGCTTTCTTCTGGGTTTGGG 678
DB 1391 GGTATATGTCCTAGAGGCTCTTCTGAGGATCTAGTTGGCTTTCTTCTGGGTTTGGG 1450
QY 679 CCAATTCAGTTCTCATGATGATGATTTCTATCATTAATGATGATGATGATGATGAT 738
DB 1451 CCAATTCAGTTCTCATGATGATGATTTCTATCATTAATGATGATGATGATGATGAT 1510
```

QY 739 TGGGACACAGAGAACTCTCTGTATTAACAATGAGGAATAGCCAGGCGGATCTCCAG 798
DB 1511 TGGGACACAGAGAACTCTCTGTATTAACAATGAGGAATAGCCAGGCGGATCTCCAG 1570
QY 799 CACCAATCTCTCAATGTTTTTCCACAGAGCTCTCCAGCCCAACCAATATGCGCTGTATG 858
DB 1571 CACCAATCTCTCAATGTTTTTCCACAGAGCTCTCCAGCCCAACCAATATGCGCTGTATG 1650
QY 859 TGTAGACATCTCTGCGGCTTCTAGCGCTGTCTCTCTCTTATGTTCTTTAATCAGATACT 918
DB 1631 TGTAGACATCTCTGCGGCTTCTAGCGCTGTCTCTCTCTTATGTTCTTTAATCAGATACT 1690
QY 919 GCGCTGGAAGCTTCTTATTTTACAGCGCTTGAAGAGCTCTTCTTGTCTAGTTGAATTAAT 978
DB 1691 GCGCTGGAAGCTTCTTATTTTACAGCGCTTGAAGAGCTCTTCTTGTCTAGTTGAATTAAT 1750
QY 979 GGTGTGTTTTCCGTAAATGAACAAATTAATTTAAAAAATGAAAAGTT 1027
DB 1751 GGTGTGTTTTCCGTAAATGAACAAATTAATTTAAAAAATGAAAAGTT 1799

RESULT 7

US-09-957-708-16
; Sequence 16, Application US/09957708
; Publication No. US20030031678A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Cafferey, Robert
; APPLICANT: Ali, Shujath
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
; TITLE OF INVENTION: Genes
; FILE REFERENCE: DEX-0239
; CURRENT APPLICATION NUMBER: US/09/957,708
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,746
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-957-708-16

Query Match 92.1%; Score 945.6; DB 3; Length 1418;
Best Local Similarity 97.8%; Pred. No. 1,9e-269;
Matches 1001; Conservative 0; Mismatches 19; Indels 4; Gaps 4;

QY 4 CAAGAGTTGGCCAAAAAATCAAGGTATTTGGTCCCGAACAAGCTTATCATTTACAGAT 63
DB 379 CAAGAGTTGGCCAAAAAATCAAGGTATTTGGTCCCGAACAAGCTTATCATTTACAGAT 438
QY 64 AAAACACTTGATGAGATGTTTCCCGCAAGCCCATATTTTCTTCTTCAATTCGAA 123
DB 439 AAAACACTTGATGAGATGTTTCCCGCAAGCCCATATTTTCTTCTTCAATTCGAA 498
QY 124 ACAAGCTCCAGAAAGCTGAAACATACCTTTGTCTTCTTGAAATTTTCCCTGATGT 183
DB 499 ACAAGCTCCAGAAAGCTGAAACATACCTTTGTCTTCTTGAAATTTTCCCTGATGT 558
QY 184 ATTAAGATACATTTGGCAAGAAAAGAGCAACAGATTTGGGATCCAGAGGGGAAAC 243
DB 559 ATTAAGATACATTTGGCAAGAAAAGAGCAACAGATTTGGGATCCAGAGGGGAAAC 618
QY 244 ACCATGAAGACTTAAGACACATACATGAAATTTAGCTGTAAACGTCGCCGAAAAGTCA 303
DB 619 ACCATGAAGACTTAAGACACATACATGAAATTTAGCTGTAAACGTCGCCGAAAAGTCA 678
QY 304 CTGGAACAAGAACACAGATGTATGCTCAGACATGAGATTAATTAATAACGGAATTGATCA 363
DB 679 CTGGAACAAGAACACAGATGTATGCTCAGACATGAGATTAATTAATAACGGAATTGATCA 738

QY 364 GAAATTTATCTTTTCTCCATTAAGAGCGATGTATCATCATGATGCCCAAGACAAATTGT 423
DB 739 GAAATTTATCTTTTCTCCATTAAGAGCGATGTATCATCATGATGCCCAAGACAAATTGT 798
QY 424 TCAAAAGATCAATATGATATACACTACTGTGAGCTCAAAACACTCTGCATATTAATG 483
DB 799 TCAAAAGATCAATATGATATACACTACTGTGAGCTCAAAACACTCTGCATATTAATG 858
QY 484 TACCTCTCTGCTCTCTCAAGAGTGTGTCTATTTTGGCATCATCACCTGCTGTCTT 543
DB 859 TACCTCTCTGCTCTCTCAAGAGTGTGTCTATTTTGGCATCATCACCTGCTGTCTT 918
QY 544 AGAAGAACGCTTTCTGCTGCAATGAGAGAAATCATTAACAGAGGTGGACAAAGAGGC 603
DB 919 AGAAGAACGCTTTCTGCTGCAATGAGAGAAATCATTAACAGAGGTGGACAAAGAGGC 978
QY 604 CATCTTTCTCTC-ATGCGTATTTGTCCTTGAAGCGCTTCTGAGATCTAGTTGGCTT 662
DB 979 CATCTTTCTCTC-ATGCGTATTTGTCCTTGAAGCGCTTCTGAGATCTAGTTGGCTT 1038
QY 663 TCTTCTGAGTTTGGGCAATTTCAATTCAGTTCTCATGTGTACTATT-CTATCATTAATGTA- 720
DB 1039 TCTTCTGAGTTTGGGCAATTTCAATTCAGTTCTCATGTGTACTATT-CTATCATTAATGTA- 1098
QY 721 TAACGTTTTCAAACCAAGTGGGACACAGAGAACTCATCTCTGTAATTAACAAAGAGAA 780
DB 1099 TAACGTTTTCAAACCAAGTGGGACACAGAGAACTCATCTCTGTAATTAACAAAGAGAA 1158
QY 781 AGCCAGCGGATCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACC 840
DB 1159 AGCCAGCGGATCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACC 1218
QY 841 AAATAGGCTGCTATATAGTGTAGACATCTGCGGCTTCTAGCTTGTCTCTTATG 900
DB 1219 AAATAGGCTGCTATATAGTGTAGACATCTGCGGCTTCTAGCTTGTCTCTTATG 1277
QY 901 TTTCTTATCAGATACATCTGCTGGAAGCTTTTCAATTTTACAGCGCTGGAAGAGTCTTCT 960
DB 1278 TTTCTTATCAGATACATCTGCTGGAAGCTTTTCAATTTTACAGCGCTGGAAGAGTCTTCT 1337
QY 961 TTGCTAGTGAATATGATGATGTTTTCGTAATAGCAAAATTAATTTAAAAAATG 1020
DB 1338 TTGCTAGTGAATATGATGATGTTTTCGTAATAGCAAAATTAATTTAAAAAAT 1397
QY 1021 AAAA 1024
DB 1398 GAAA 1401

RESULT 8
US-09-925-300-58
; Sequence 58, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruden,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (135)

OY	59	CAGATTAACAACCTTGATGTCAGATGTTTTCCCCCAACCCGACATATTTTTCTTCCTTCAATTG	118
Db	160	CTGATTAACAACCTTGATGTCAGATGTTTTCCCCCAAGCCACATG-TTTTCTTCCTTCRATTTG	218
OY	119	CTGAAACCAAGCCTCAGAAAGCTGGAAACATACCTTTGTCTTCTTGAGAAATTTTTTCCCTG	178
Db	219	CTGAAACCAAACTCCAGAAAGCTGGAAACATACCTTTGTCTTCTTGAGAAATTTTTTCCCMG	278
OY	179	ATGTATTATAATATCATTTGGCAAGAAAAGAAAGGACAACGATTCCTGGGATCCGAGAGG	238
Db	279	ATRTATTAAAGTATCATTTGSSAAGAAAAGAAAGGACAACGATTCCTGGGATCCGAGAGG	338
OY	239	GG--AAACCACTGAAGACTTAAAGACACATACATGAATTAATTTAGCTGGTTTAAACGGTCCAGA	296
Db	339	GGGAACACCATGGAAGACTTAAAGACACATACATGAATTAATTTAGCTGGTTTAAACGGTCCAGA	398
OY	297	AAAGTCACTGGAACAAAGAACACAGATGTATGCT-CAGACATGAG-AATATATTAAGAACGA	354
Db	399	AAAGTCACTGGAACAAAGAACACAGATGTATGCTTCCATGACATGAGAAATATTAAGAACGR	458
OY	355	GTTGATCAAGAAATATATCTTTTCCTCCATTAAGACGAGATGTATCATACATGGATCCCAAA	414
Db	459	GTTGATCAAGAAATATATCTTTTCCTCCATTAAGACAGATGTATCATACATGGATCCCAAA	518
OY	415	GACAAATTGTTCAAAAGATGCAAAATGATACATGCTGTGACGCTCAACAAACCTCTGTGA	474
Db	519	GACAAATTGTTCAAAAGATGCAAAATGATACACTACGCTGACGCTCAACAAACCTCTGTGA	578
OY	475	TATTACATGTAACCTTCCTGCTGCTCAAGAGTGGTGTATTTTGGCATGATCATGACCTGG	534
Db	579	TATTACATGTAACCTTCCTGCTGCTCAAGAGTGGTGTATTTTGGCATGATCATGACCTGG	638
OY	535	TGTCGTCTTAAGAAAGACGGCTTTTCTGCTGCATGAGAGAAATCATTAACAGCGGTGGCA	594
Db	639	TGTCGTCTTAAGAAAGACGGCTTTTCTGCTGCATGAGAGAAATCATTAACAGCGGTGGCA	698
OY	595	CAAGGAGGCAATCTTTTCTCTCATCGGTTATGTCTCTAGAAAGCGTCTTCTGAGATCTAG	654
Db	699	CAAGGAGGCAATCTTTTCTCTCATCGGTTATGTCTCTAGAAAGCGTCTTCTGAGATCTAG	758
OY	655	TTGGGCTTTCTTTCTGGGTTGGGTCATTTGAGTTCTGATGATGATGATCTATCTATCAAT	714
Db	759	TTGGGCTTTCTTTCTGGGTTGGGTCATTTGAGTTCTGATGATGATGATCTATCTATCAAT	818
OY	715	ATTGTATATAAGGTTTTCAAACCAAGTGGGCAACAAGAAACCTCACTCTGTATAAACAATG	774
Db	819	ATTGTATATAAGGTTTTCAAACCAAGTGGGCAACAAGAAACCTCACTGTATAAACAATG	878
OY	775	AGGAATATACCAACGGGATCTCCAGACAACAAATCTCCATGTTTTTCAACAGCTCCACAGC	834
Db	879	AGGAATATACCAACGGGATCTCCAGACAACAAATCTCCATGTTTTTCAACAGCTCCACAGC	938
OY	835	CAACCCAATATAGCGCTGCTATAGTGTAGACATCTGCGGGCTTTCAGCCTTGTCCCTCTC	894
Db	939	CAACCCAATATAGCGCTGCTATAGTGTAGACATCTGCGGGCTTTCAGCCTTGTCCCTCTC	998
OY	895	TTAGTGTTCTTTAATCAAGTATATCTGCTGGAGCCTTTTCACTTTTACAAGCCTCGAAGCAG	954
Db	999	TTAGTGTTCTTTAATCAAGTATATCTGCGAAGCCTTTTCACTTTTACAAGCCTCGAAGCAG	1058

```

; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 317
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-317

Query Match      87.8%; Score 902.2; DB 3; Length 1421;
Best Local Similarity 92.8%; Pred. No. 1,5e-256;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

QY      20  AATCAAGTATTGGTCCCGGACAAAGCTTATCATTTACAGATAAACCACTTGATGAG 79
DB      369  ATAAAGAACTCTTTGGCAGTGGAGCAACACTTGTTGCACAGATTAACCACTTATGCGAG 428
QY      80  ATGTTTCCCGCAAGCCCACTATTTTCTTCCTTCAATTTGCTGAAACAAAGCTCCAGAAG 139
DB      429  ATGTTTCCCGCAAGCCCACTATTTTCTTCCTTCAATTTGCTGAAACAAACCTCCAGAAG 488
QY      140  CTGGAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAGATACATTGGC 199
DB      489  CTGGAACATACCTTTGTCTTCTTGAAGAAATTTTCCAGATATTATTAAAGTATACATTGGC 548
QY      200  AAGAAAAAGAGCAACAGATTTCTGGATCCAGAGAGGGAAACCACTAAGACTAAG 259
DB      549  AAGAAAAAGAGCAACAGATTTCTGGATCCAGAGAGGGAAACCACTAAGACTAAG 608
QY      260  ACACATACATGAATTTAGTCTGGTTAAGCGTGCAGAAAAGTCACTGACAAAGAACCA 319
DB      609  ACACATACATGAATTTAGTCTGGTTAAGCGTGCAGAAAAGTCACTGACAAAGAACCA 668

```

D6	669	GATGTATCCGTCAAGCATTGAAATTAATAAACGGAATTGATCAAGAAATTATCTTCTCC	728
QY	380	CAATTAAGAC-----GG	391
D6	729	CAATTAAGACAGATGTGTACCCACAGTGGATCCCAAGACAGTTATTCAAAAGATGCAATG	788
QY	392	ATGTCATCACAAATGGATATCCCAAGACAAATTGTTCAAAAGATGCAATGCAATACATACG	451
D6	789	ATGTCATCACAAATGGATATCCCAAGACAAATTGTTCAAAAGATGCAATACATACATACG	848
QY	452	TGCAGCTCACAAACACCTCTGCATATTACATGTACCTCTCCGCTCTCCCAAGATGTGG	511
D6	849	TGCAGCTCACAAACACCTCTGCATATTACATGTACCTCTCCGCTCTCCCAAGATGTGG	908
QY	512	TCATATTTGCCATCATCATCCTGCTGTCTGCTTAGAAGAAACGCTTTCTGCTGCAATGAG	571
D6	909	TCATATTTGCCATCATCATCCTGCTGTCTGCTTAGAAGAAACGCTTTCTGCTGCAATGAG	968
QY	572	AGAAATATATAACAGACGCTGTGCACAGAGAGCCATCTTTTCTCATCGGTATATGTCCT	631
D6	969	AGAAATATATAACAGACGCTGTGCACAGAGAGCCATCTTTTCTCATCGGTATATGTCCT	1028
QY	632	AGAAAGCTTTCTGAGGATCTAGTGGGCTTTCTTTCTGGGCTTTGGGCCATTTCAGTTCT	691
D6	1029	AGAAAGCTTTCTGAGGATCTAGTGGGCTTTCTTTCTGGGCTTTGGGCCATTTCAGTTCT	1088
QY	692	CATGTGTACTATTCTATCATTTATTTGTATATACGATTTTCAACACAGTGGGACACAGAG	751
D6	1089	CATGTGTACTATTCTATCATTTATTTGTATATACGATTTTCAACACAGTGGGACACAGAG	1148
QY	752	AACTCTACTCTGTATATAACAATGAGAAATAGCCACGGCGATCTCCAGACCAATCTCTCC	811
D6	1149	AACTCTACTCTGTATATAACAATGAGAAATAGCCATGCGATCTCCAGACCAATCTCTCC	1208
QY	812	ATGTTTTCAACAGCTCCTCCAGCCAAACCCAAATATAGGCGCTGTATAGTAGACATCTCG	871
D6	1209	ATGTTTTCAACAGCTCCTCCAGCCAAACCCAAATATAGGCGCTGTATAGTAGACATCTCG	1268
QY	872	CGGCTTTCAGCTTGTCCCTCTCTTTAGTGTCTTTATATCAGATATATGCTCTGGAAGCTT	931
D6	1269	CGGCTTTCAGCTTGTCCCTCTCTTTAGTGTCTTTATATCAGATATATGCTCTGGAAGCTT	1328
QY	932	TCATTTTACACGCGCTTAAGCAGCTCTTTCGTAAGTATATATGCTGTGTGTTTTCC	991
D6	1329	TCATTTTACACGCGCTTAAGCAGCTCTTTCGTAAGTATATATGCTGTGTGTTTTCC	1388
QY	992	CTAATTAAGCAAAATTAATTTTAAAAAATGAAAA	1024
D6	1389	GTAATTAAGCAAAATTAATTTTAAAAAATGAAAA	1421

```

: PRIOR APPLICATION NUMBER: US/09/964,824
: PRIOR FILING DATE: 2001-09-27
: PRIOR APPLICATION NUMBER: US/09/967,766
: PRIOR FILING DATE: 2001-09-28
: PRIOR APPLICATION NUMBER: US/09/968,007
: PRIOR FILING DATE: 2001-10-02
: PRIOR APPLICATION NUMBER: US/09/969,347
: PRIOR FILING DATE: 2001-10-02
: PRIOR APPLICATION NUMBER: US/09/969,708
: PRIOR FILING DATE: 2001-10-03
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 8447
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3344
: LENGTH: 1421
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-843-641A-3344

Query Match      87.8%; Score 902.2; DB 9; Length 1421;
Best Local Similarity 92.8%; Pred. No. 1.5e-256;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1

QY 20 AATCAAGATATTGGTCCGGAAACAAGCTTATCATACAGATTAACAACCTTGATGAG 79
Db 369 ATTAAGAACTTTGGCAGTGGAAACAACCTTTGTGTACAGATTAACAACCTTGATGAG 428
QY 80 ATGTTTCCCCCAGCCCACTATTTTCTTCCTTCAATTGCTGAACAAGCTCCAGAAG 139
Db 429 ATGTTTCCCCCAGCCCACTATTTTCTTCCTTCAATTGCTGAACAAGCTCCAGAAG 488
QY 140 CTGGAACATCACTTTGTCTTCTTGAGAAATTTTCCCTGAGATTTAATGATCACTTGGC 199
Db 489 CTGGAACATCACTTTGTCTTCTTGAGAAATTTTCCCGAATTTATTAAGATCACTTGGC 548
QY 200 AAGAAAAGAGAAGCAACAAGATTCCTGGGATCCCAAGAGAGGGAACACATGAAGCTAAAG 259
Db 549 AAGAAAAGAGAAGCAACAAGATTCCTGGGATCCCAAGAGAGGGAACACATGAAGCTAAAG 608
QY 260 ACACATACATGAATTTAGCTGTGTTAACGATGCCAGAAAGTCACTGACAAAGAACACA 319
Db 609 ACACATACATGAATTTAGCTGTGTTAACGATGCCAGAAAGTCACTGACAAAGAACACA 668
QY 320 GATGATTCGTACAGACATAGAAATTAATAAAAAGAGTTGATCAAGAAATTAATCTTCTCC 379
Db 669 GATGATTCGTACAGACATAGAAATTAATAAAAAGAGTTGATCAAGAAATTAATCTTCTCC 728
QY 380 CAATTAACAC-----GG 391
Db 729 CAATTAACACAGATGTCAACAAGTGTATCCCAAGACAGTTATTCAAAAGATGCAAATG 788
QY 392 ATGTCAATCAATGGATCCCAAGACAAATTTGTTCAAAAGATGCAATGATATCACTACTGC 451
Db 789 ATGTCAATCAATGGATCCCAAGACAAATTTGTTCAAAAGATGCAATGATATCACTACTGC 848
QY 452 TGCAGCTCAAAACACCTCTGATATTAATGATACCTCTCTGCTCTCAAGAGTGTG 511
Db 849 TGCAGCTCAAAACACCTCTGATATTAATGATACCTCTCTGCTCTCAAGAGTGTG 908
QY 512 TCTATTTTGCATATCACTCTGCTGTCTGCTTTAAGAAAGACGGCTTTGCTGCAATGAG 571
Db 909 TCTATTTTGCATATCACTCTGCTGTCTGCTTTAAGAAAGACGGCTTTGCTGCAATGAG 968
QY 572 AGAATCAATTAACAGACGGTGGCAAGAGGCACTTTTCTCTCATCGGTTATTTGCTCT 631
Db 969 AGAATCAATTAACAGACGGTGGCAAGAGGCACTTTTCTCTCATCGGTTATTTGCTCT 1022
QY 632 AGAAGCGCTCTCTGAGATCTAGTTGAGGCTTCTTCTTGAGTTTGGGCCATTTCAAGTTC 691
Db 1029 AGAAGCGCTCTCTGAGATCTAGTTGAGGCTTCTTCTTGAGTTTGGGCCATTTCAAGTTC 1088
QY 692 CATGTGTACTATTCTATCATTAATTGATTAACGGTTTCAAAACAGTGGGACACAGAG 751

```

Db 1089 CATGTGTACTATTTCTATCATTTATGTATATGTTTCAAAACAGTGGGACACAGAG 1148
Qy 752 AACCTCACTGTATTAACAATAGGAATAGCCAGCGGAGTCTCCAGCAACATCTCTCC 811
Db 1149 AACCTCACTGTATTAACAATAGGAATAGCCAGTGGGAGTCTCCAGCAACATCTCTCC 1208
Qy 812 ATGTTTCCACAGCTCTCCAGCAACCAAAATAGCCGCTATATAGTGAACATCTG 871
Db 1209 ATGTTTCCACAGCTCTCCAGCAACCAAAATAGCCGCTATATAGTGAACAGCTG 1268
Qy 872 CGGCTTACGCTTGTCTCTCTTATGTTTCTTTAATCAGATTAAGTCTGGAAGCTT 931
Db 1269 CGGCTTACGCTTGTCTCTCTTATGTTTCTTTAATCAGATTAAGTCTGGAAGCTT 1328
Qy 932 TCATTTTACAGCGCTTGAAGAGTCTCTTCTAGTTGAATTAATGTTGTTTCTCC 991
Db 1329 TCATTTTACAGCGCTTGAAGAGTCTCTTCTAGTTGAATTAATGTTGTTTCTCC 1388
Qy 992 GTAAATAGCAAAATTAATTTAAAAAATGAAAA 1024
Db 1389 GTAAATAGCAAAATTAATTTAAAAAATGAAAA 1421

RESULT 11
US-09-960-706-676
; Sequence 676, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 676
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M30894
US-09-960-706-676

Query Match 81.6%; Score 838; DB 3; Length 1586;
Best Local Similarity 88.4%; Pred. No. 1.8e-237;
Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;
Qy 20 AAATCAAGTATTTGTCCTCCGGAACAAGCTTATCATTAAGATTAACAAGTGTGATCAG 79
Db 499 ATAAGAACTCTTGGCAGTGGAGCAACAATTGTGTCAAGATTAACAAGTGTGATCAG 558
Qy 80 ATGTTTCCCAAGCCCACTATTTTCTCTCTCAATTTGCTGAACAAGTCTCCAGAAG 139
Db 559 ATGTTTCCCAAGCCCACTATTTTCTCTCTCAATTTGCTGAACAAGTCTCCAGAAG 618
Qy 140 CTGGAACATGCTTGTCTTCTGAGAATTTTCCCGATGTAATTAAGATACATTTGGC 199
Db 619 CTGGAACATGCTTGTCTTCTGAGAATTTTCCCGATGTAATTAAGATACATTTGGC 678
Qy 200 AAGAAAGAGAGCAACAGATTTCTGGATCTCCAGAGGGGAAACCATGAAGAATAAG 259
Db 679 AAGAAAGAGAGCAACAGATTTCTGGATCTCCAGAGGGGAAACCATGAAGAATAAG 738
Qy 260 ACAATATACATGAATTTTACGCTGTTAAACGCTGCCAGAAAGTCACTGACAAAGAACCA 319
Db 739 ACAATATACATGAATTTTACGCTGTTAAACGCTGCCAGAAAGTCACTGACAAAGAACCA 798
Qy 320 GATGTATGTCAGACATGAGAAATAAATAAAGGAGTTGATCAAGAAATATCTTCTC 379

Db 799 GATGTATGTCAGACATGAGAAATAAATAAAGGAGTTGATCAAGAAATATCTTCTC 858
Qy 380 CATTAAGAC----- 389
Db 859 CAATTAAGACAGATGTACACACAGTGAATCCAAAGACAGTATTCAAAAAGATGAATG 918
Qy 390 -----GATGTATCAACAA 403
Db 919 ATGTCAACACAGTGAATCCAAATTAACATTTATCAAGAGATGAAGATGATGATCAAC 978
Qy 404 TGGATCCAAAGACAAATTTGTTCAAAAGATGAAGATGATGATGATGATGATGATGAT 463
Db 979 TGGATCCAAAGACAAATTTGTTCAAAAGATGAAGATGATGATGATGATGATGATGAT 1038
Qy 464 ACACCTTGTGATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 523
Db 1039 ACACCTTGTGATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
Qy 524 TCATCACTGCTGTCTGTAGAGAGAGGCTTCTGTGCAATGAGAGAGAGAGAGAGAGAG 583
Db 1099 TCATCACTGCTGTCTGTAGAGAGAGGCTTCTGTGCAATGAGAGAGAGAGAGAGAGAG 1158
Qy 584 AGACGTTGCAACAGAGAGGCAATCTTTCTCATGCGTATTTGCTTCTGAGAGAGAGAG 643
Db 1159 AGACGTTGCAACAGAGAGGCAATCTTTCTCATGCGTATTTGCTTCTGAGAGAGAGAG 1218
Qy 644 TGAAGATTAAGTGGGCTTCTTCTGAGTGGGCTTCTGAGTGGGCTTCTGAGTGGGCTT 703
Db 1219 TGAAGATTAAGTGGGCTTCTTCTGAGTGGGCTTCTGAGTGGGCTTCTGAGTGGGCT 1278
Qy 704 ATTCTATCATTTATTTATTAACAGTTTCAACAGTGGGCAACAGAGAGAGAGAGAGAGAG 763
Db 1279 ATTCTATCATTTATTTATTAACAGTTTCAACAGTGGGCAACAGAGAGAGAGAGAGAGAG 1338
Qy 764 TAATTAACAATGAGAAATAGCAGGCGATCTCCAGACCAATCTCTCATGTTTTCACA 823
Db 1339 TAATTAACAATGAGAAATAGCAGGCGATCTCCAGACCAATCTCTCATGTTTTCACA 1398
Qy 824 GCTCTCCAGCCCAACCAATAGCGCTGTATAGTGAACATCTCTGCGGCTTCTAGCC 883
Db 1399 GCTCTCCAGCCCAACCAATAGCGCTGTATAGTGAACATCTCTGCGGCTTCTAGCC 1458
Qy 884 TTGCTCTCTCTAGTGTCTTATTAACAGTGAATGCTGGAAGGCTTCTATTTTACAG 943
Db 1459 TTGCTCTCTCTAGTGTCTTATTAACAGTGAATGCTGGAAGGCTTCTATTTTACAG 1518
Qy 944 CCTGAAGCAGTCTTCTTGTGATTAATTAATGATGATGATGATGATGATGATGATGAT 1003
Db 1519 CCTGAAGCAGTCTTCTTGTGATTAATTAATGATGATGATGATGATGATGATGATGAT 1578
Qy 1004 ATAAATTT 1011
Db 1579 ATAAATTT 1586

RESULT 12
US-09-873-319-424
; Sequence 424, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Yamamoto, Jun
; APPLICANT: Waga, Iwao
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; PRIOR FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 755
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 424
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M30894
US-09-873-319-424

Query Match 81.6%; Score 838; DB 3; Length 1586;
Best Local Similarity 88.4%; Pred. No. 1.8e-237;
Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;

20 AAATCAAGTATTGTCCTCCGGAACAAAGCTTATCATACAGATTAACCACTTGATGAG 79
499 ATAAAGAACTCTTGGCAGTGAACAAACATTGTTGTACAGATTAACCACTTGATGAG 558
80 ATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCTGAACCAAGCTCCAGAAAG 139
559 ATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCTGAACCAAGCTCCAGAAAG 618
140 CTGGAACATACCTTGTCTCTGAGAAATTTTCCCGATGTTAATAGATGATGAGC 199
619 CTGGAACATATCTTGTCTCTGAGAAATTTTCCCGATGTTAATAGATGATGAGC 678
200 AAGAAAGAAAGACACAGATTCGGATCCAGAGGGGAAACCATGAAGACTAAGC 259
679 AAGAAAGAAAGACACAGATTCGGATCCAGAGGGGAAACCATGAAGACTAAGC 738
260 ACACATACATGAATTTTACTGCTTACGGTCCAGAAAGTCACTGACAAAGACACA 319
739 ACACATACATGAATTTTACTGCTTACGGTCCAGAAAGTCACTGACAAAGACACA 798
320 GATGATGCTGACACATGAGAAATTAATAACGGAGTTGATGAAGAAATTTCTTCTC 379
799 GATGATGCTGACACATGAGAAATTAATAACGGAGTTGATGAAGAAATTTCTTCTC 858
380 CAATTAAGAC----- 389
859 CAATTAAGACAGATGTACCAACAGTATCCCAAGACAGTTATTCAAAAGATCAAAATG 918
390 -----GATGTATCAACAA 403
919 ATGTACACACAGATGTATCCAAATTAATTAAGATGACAAATGATGTATCAACAA 978
404 TGGATCCCAAGACAAATTTGTTCAAAAGATGACAACTACTGTGAGCTCAACAA 463
979 TGGATCCCAAGACAAATTTGTTCAAAAGATGACAACTACTGTGAGCTCAACAA 1038
464 ACACCTGACATATTAACATGTAACCTCTCTGCTCTCAAGAGTGTGTTATTGTTGCA 523
1039 ACACCTGACATATTAACATGTAACCTCTCTGCTCTCAAGAGTGTGTTATTGTTGCA 1098
524 TCATCACTGCTGTCTGTTAGAGACGGCTTTCTGCTGCAATGAGAGAAATCAATAC 583
1099 TCATCACTGCTGTCTGTTAGAGACGGCTTTCTGCTGCAATGAGAGAAATCAATAC 1158
584 AGACGGTGGCAACAGAGGCAATCTTTCCTCATGCGTTATTTGCCCTGAAGGCTCTTC 643
1159 AGACGGTGGCAACAGAGGCAATCTTTCCTCATGCGTTATTTGCCCTGAAGGCTCTTC 1218
644 TGAAGATCTAGTGGCTTTCTTCTGAGTTTGGGCAATTCAGTTCTCATGTGTACT 703
1219 TGAAGATCTAGTGGCTTTCTTCTGAGTTTGGGCAATTCAGTTCTCATGTGTACT 1278
704 ATTCTATCATTAATGTAATGCGTTTCAACAGTGGGCAACAGAGAACTCACTG 763
1279 ATTCTATCATTAATGTAATGCGTTTCAACAGTGGGCAACAGAGAACTCACTG 1338
764 TAAATTAACAATGAGAAATGCGGCAATTCACAGACCAATCTCTCATGTTTCCACA 823
1339 TAAATTAACAATGAGAAATGCGGCAATTCACAGACCAATCTCTCATGTTTCCACA 1398

824 GCTCTCCACCAACCCAAATAGCGCTGCTATAGTGTAGACATCTGCGGCTTACGC 883
1399 GCTCTCCACCAACCCAAATAGCGCTGCTATAGTGTAGACAGCCTGCGGCTTACGC 1458
884 TTGTCCCTCTTATAGTGTCTTTTATTCAGATTAACCTGCTGAAAGCCTTTCATTTACAG 943
1459 TTGTCCCTCTTATAGTGTCTTTTATTCAGATTAACCTGCTGAAAGCCTTTCATTTACAG 1518
944 CCTGAAGCAGTCTTCTTCTGATGAATTAATGATGATGTTTTCGTAATTAAGCAAA 1003
1519 CCTGAAGCAGTCTTCTTCTGATGAATTAATGATGATGTTTTCGTAATTAAGCAAA 1578
1004 ATAAATTT 1011
1579 ATAAATTT 1586

RESULT 13
US-10-847-918-14
; Sequence 14, Application US/10847918
; Publication No. US20050119210A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Be, Xiaobing
; APPLICANT: Liu, Wei
; APPLICANT: Slonim, Donna
; APPLICANT: Howes, Steve
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating Cancers
; FILE REFERENCE: 031896-026000 (JMI01264)
; CURRENT APPLICATION NUMBER: US/10/847,918
; PRIOR FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US 60/471,729
; NUMBER OF SEQ ID NOS: 14937
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 14
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-847-918-14

Query Match 81.6%; Score 838; DB 9; Length 1586;
Best Local Similarity 88.4%; Pred. No. 1.8e-237;
Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;

20 AAATCAAGTATTGTCCTCCGGAACAAAGCTTATCATACAGATTAACCACTTGATGAG 79
499 ATAAAGAACTCTTGGCAGTGAACAAACATTGTTGTACAGATTAACCACTTGATGAG 558
80 ATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCTGAACCAAGCTCCAGAAAG 139
559 ATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCTGAACCAAGCTCCAGAAAG 618
140 CTGGAACATACCTTGTCTTGAAGAAATTTTCCCTGATGTTAATAGATGATGAGC 199
619 CTGGAACATACCTTGTCTTGAAGAAATTTTCCCTGATGTTAATAGATGATGAGC 678
200 AAGAAAGAAAGACAAACATCTGGGATCCCAAGAGGGGAAACCATGAAGACTAAGC 259
679 AAGAAAGAAAGACAAACATCTGGGATCCCAAGAGGGGAAACCATGAAGACTAAGC 738
260 ACACATACATGAATTTTACTGCTTACGGTCCAGAAAGTCACTGACAAAGACACA 319
739 ACACATACATGAATTTTACTGCTTACGGTCCAGAAAGTCACTGACAAAGACACA 798
320 GATGATGCTGACACATGAGAAATTAATAACGGAGTTGATGAAGAAATTTCTTCTC 379
799 GATGATGCTGACACATGAGAAATTAATAACGGAGTTGATGAAGAAATTTCTTCTC 858
380 CAATTAAGAC----- 389
859 CAATTAAGACAGATGTACCAACAGTATCCCAAGACAGTTATTCAAAAGATCAAAATG 918

RESULT 15
US-10-357-930-24142/c
; Sequence 24142, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endegre, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BEN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 24142
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-24142

Query Match 77.5%; Score 795.8; DB 8; Length 1162;
Best Local Similarity 99.7%; Pred. No. 5.3e-225;
Matches 797; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 229 TCCGAGGAGGGAACCATGAAGCTAAGACACATCATGAATTTAGCTGGTTACG 288
DB 1162 TCCGAGGAGGGAACCATGAAGCTAAGACACATCATGAATTTAGCTGGTTACG 1103
QY 289 GTGCAGAAAGTCACTGCAAGAAACACAGATGTATCTCAGACATGAGATATATAA 348
DB 1102 GTGCAGAAAGTCACTGCAAGAAACACAGATGTATCTCAGACATGAGATATATAA 1043
QY 349 AACGAGTTGATCAAGAAATATCTTTCCTCAATAAAGCGATGTATCAATGAT 408
DB 1042 AACGAGTTGATCAAGAAATATCTTTCCTCAATAAAGCGATGTATCAATGAT 983
QY 409 CCCAAAGCAATTTGTTCAAAAGATGCAATATACATCTGCTGCACTCAACAAACCC 468
DB 982 CCCAAAGCAATTTGTTCAAAAGATGCAATATACATCTGCTGCACTCAACAAACCC 923
QY 469 TCTGATATTACATATACCTCTCTGCTCTCAAGAGTGTGTATTTTGGCATCATC 528
DB 922 TCTGATATTACATATACCTCTCTGCTCTCAAGAGTGTGTATTTTGGCATCATC 863
QY 529 ACCTGCTGTCTGTAGAAAGCGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGC 588
DB 862 ACCTGCTGTCTGTAGAAAGCGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGC 803
QY 589 GTGGACAAGAGAGGCACTTTTCTCTCATCGTTATTTGCTCTAGAACGCTTTTGAGG 648
DB 802 GTGGACAAGAGAGGCACTTTTCTCTCATCGTTATTTGCTCTAGAACGCTTTTGAGG 743
QY 649 ATCTAGTGGGCTTTCTTCTGAGTTTGGGCAATTTGAGTTCTCATGTGTACTATTCT 708
DB 742 ATCTAGTGGGCTTTCTTCTGAGTTTGGGCAATTTGAGTTCTCATGTGTACTATTCT 683

QY 709 ATCATTTTGTATTAACGGTTTTCAAACCAAGTGGGACACAGAGAACTCACTGTATATA 768
DB 682 ATCATTTTGTATTAACGGTTTTCAAACCAAGTGGGACACAGAGAACTCACTGTATATA 623
QY 769 ACAATGAGATAGCCAGCGGATCTCCAGACCAATCTCTCCATGTTTCCACAGCTCC 828
DB 622 ACAATGAGATAGCCAGCGGATCTCCAGACCAATCTCTCCATGTTTCCACAGCTCC 563
QY 829 TCCAGCCAAACCAATAGCGCTGCTATAGTATAGACATCTGCGGCTTTAGCCTTGTG 888
DB 562 TCCAGCCAAACCAATAGCGCTGCTATAGTATAGACATCTGCGGCTTTAGCCTTGTG 503
QY 889 CCTCTTATAGTCTTTTATCAATTAACATGCTGGAAGCTTTGATTTTACAGCCCTG 948
DB 502 CCTCTTATAGTCTTTTATCAATTAACATGCTGGAAGCTTTGATTTTACAGCCCTG 443
QY 949 AAGCAGCTTCTTGTAGTTGAATTAATGTGTGTGTTTTCGTAATAGCAAAATATA 1008
DB 442 AAGCAGCTTCTTGTAGTTGAATTAATGTGTGTGTTTTCGTAATAGCAAAATATA 383
QY 1009 TTTAAAAAATGAAAAGTT 1027
DB 382 TTTAAAAAATGAAAAGTT 364

Search completed: December 11, 2005, 01:30:51
Job time : 895 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 20:30:37 ; Search time 230 Seconds
(without alignments)
7937.199 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027

Sequence: 1 GGGCAAGAGCTGGGCAAGAAAAA.....ATTCAAAAGATGAAAGCT 1027

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata1/ina/1.COMB.seq: *
2: /cgn2_6/prodata1/ina/5.COMB.seq: *
3: /cgn2_6/prodata1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata1/ina/H.COMB.seq: *
6: /cgn2_6/prodata1/ina/PCTUS.COMB.seq: *
7: /cgn2_6/prodata1/ina/PP.COMB.seq: *
8: /cgn2_6/prodata1/ina/RE.COMB.seq: *
9: /cgn2_6/prodata1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	516.4	50.3	1080	9	5260223-3 Patent No. 5260223
2	449.2	43.7	825	2	US-08-256-964A-18 Sequence 18, Appl
3	444	43.2	511	3	US-09-020-956-72 Sequence 72, Appl
4	444	43.2	511	3	US-09-030-607-72 Sequence 72, Appl
5	444	43.2	511	3	US-09-439-313-72 Sequence 72, Appl
6	444	43.2	511	3	US-09-352-616A-72 Sequence 72, Appl
7	444	43.2	511	3	US-09-232-149A-72 Sequence 72, Appl
8	444	43.2	511	3	US-09-159-812-72 Sequence 72, Appl
9	444	43.2	511	3	US-09-636-215-72 Sequence 72, Appl
10	444	43.2	511	3	US-09-166A-72 Sequence 72, Appl
11	444	43.2	511	3	US-09-115-453-72 Sequence 72, Appl
12	444	43.2	511	3	US-09-688-489-72 Sequence 72, Appl
13	444	43.2	511	3	US-09-679-426-72 Sequence 72, Appl
14	444	43.2	511	3	US-09-759-143-72 Sequence 72, Appl
15	444	43.2	511	3	US-09-651-236-72 Sequence 72, Appl
16	444	43.2	511	3	US-09-030-606-72 Sequence 72, Appl
17	444	43.2	511	3	US-09-657-279-72 Sequence 72, Appl
18	444	43.2	511	3	US-10-012-896-72 Sequence 72, Appl
19	407	39.6	467	3	US-09-020-956-75 Sequence 75, Appl
20	407	39.6	467	3	US-09-030-607-75 Sequence 75, Appl
21	407	39.6	467	3	US-09-439-313-75 Sequence 75, Appl
22	407	39.6	467	3	US-09-352-616A-75 Sequence 75, Appl
23	407	39.6	467	3	US-09-232-149A-75 Sequence 75, Appl
24	407	39.6	467	3	US-09-159-812-75 Sequence 75, Appl

25	407	39.6	467	3	US-09-636-215-75 Sequence 75, Appl
26	407	39.6	467	3	US-09-685-166A-75 Sequence 75, Appl
27	407	39.6	467	3	US-09-115-453-75 Sequence 75, Appl
28	407	39.6	467	3	US-09-688-489-75 Sequence 75, Appl
29	407	39.6	467	3	US-09-679-426-75 Sequence 75, Appl
30	407	39.6	467	3	US-09-759-143-75 Sequence 75, Appl
31	407	39.6	467	3	US-09-651-236-75 Sequence 75, Appl
32	407	39.6	467	3	US-09-030-606-75 Sequence 75, Appl
33	407	39.6	467	3	US-09-657-279-75 Sequence 75, Appl
34	407	39.6	467	3	US-10-012-896-75 Sequence 75, Appl
35	404.4	39.4	426	3	US-09-513-999C-679 Sequence 679, App
36	250.6	24.4	281	3	US-09-439-313-433 Sequence 433, App
37	250.6	24.4	281	3	US-09-352-616A-433 Sequence 433, App
38	250.6	24.4	281	3	US-09-636-215-433 Sequence 433, App
39	250.6	24.4	281	3	US-09-685-166A-433 Sequence 433, App
40	250.6	24.4	281	3	US-09-679-426-433 Sequence 433, App
41	250.6	24.4	281	3	US-09-759-143-433 Sequence 433, App
42	250.6	24.4	281	3	US-09-651-236-433 Sequence 433, App
43	250.6	24.4	281	3	US-09-657-279-433 Sequence 433, App
44	250.6	24.4	281	3	US-10-012-896-433 Sequence 433, App
45	247.6	24.1	460	3	US-09-513-999C-8197 Sequence 8197, Ap

ALIGNMENTS

RESULT 1									
5260223-3									
Patent No. 5260223									
; APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN, JOHN G.; IP, STEPHEN H.; KRANGEL, MICHAEL S.									
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA ; T CELL RECEPTOR									
; NUMBER OF SEQUENCES: 4									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/07/187,698									
; FILING DATE: 29-APR-1988									
; PRIOR APPLICATION DATA:									
; APPLICATION NUMBER: 115,256									
; FILING DATE: 29-OCT-1987									
; APPLICATION NUMBER: 16,252									
; FILING DATE: 19-FEB-1987									
; APPLICATION NUMBER: 882,100									
; FILING DATE: 03-JUL-1986									
; SEQ ID NO:3:									
; LENGTH: 1080									
5260223-3									
Query Match									
Best Local Similarity 50.3%; Score 516.4; DB 9; Length 1080;									
Best Local Similarity 88.9%; Pred. No. 8.4e-150;									
Matches 590; Conservative 0; Mismatches 26; Indels 48; Gaps 1;									
QY	25	AAAGTATTTGTC	CCGAA	CAAGCTT	ATCAGAT	TAACAACT	TGATG	AGATGTT	84
DB	403	AAACTTTTGGAG	TGAACA	CACTTGT	GCAGAT	TAACAACT	TGATG	CAATGT	462
QY	85	TTCCCAAGCCCA	CTATTTT	CTTCTT	CAATTG	CTGAA	CAAAAG	CTCCGA	144
DB	463	TTCCCAAGCCCA	CTATTTT	CTTCTT	CAATTG	CTGAA	CAAAAG	CTCCGA	522
QY	145	ACATACCTTGT	CTTCTG	GAATTTT	CCCTG	ATGTTAT	TAA	AGATCTG	204
DB	523	ACATACCTTGT	CTTCTT	GAATTTT	CCCTG	ATGTTAT	TAA	AGATCTG	582
QY	205	AAAGAGGACA	CGATTTG	GATCC	AGAGG	GGAA	CA	CCATGA	264
DB	583	AAAGAGGACA	CGATTTG	GATCC	AGAGG	GGAA	CA	CCATGA	642
QY	265	TACATGAATTT	AGCTG	TTAAG	CGTCC	CAAA	AAATC	CTGCA	324
DB	643	TACATGAATTT	AGCTG	TTAAG	CGTCC	CAAA	AAATC	CTGCA	702
QY	325	ATGTCAGACAT	GAGAT	ATATATA	AAACG	AGTTG	ATCA	AGAAAT	384


```

; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yugu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kaios, Michael
; APPLICANT: Rafter, Gary
; APPLICANT: Rafter, Mark
; APPLICANT: Soik, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(511)
; OTHER INFORMATION: n = A,T,C or G
US-09-439-313-72

```

```

Query Match          43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1.7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

Qy 492 CCTGCTCTCAAGAGTGTGTCTATTTTGGCCATCATCCTGCTGTCTTGAAGAAC 551
Db 506 CNGGCTCTCAAGAGGAGTGTATTTTNGCCATCATCCTGNGTTGTTGAAGAACG 447
Qy 552 GCGTTTGTGTCGATGAGAGAAATCATACAGCGGTGACAGAGGCGCATCTTTT 611
Db 446 GGTITNMGTCGATGAGAGAAATCATACAGCGGCGACAGAGGCGCATCTTTT 387
Qy 612 CCTCATCGTTATTTGCTCCCTAGAAAGCGCTTCTGAGATCTAGTTGGCTTTCTTTG 671
Db 386 CTTATCGTTATTTGCTCCCTAGAAAGCGTNTCTGAGATCTAGTTGGCTTTCTTTG 327
Qy 672 GTTTGGGCAATTTCAATTTCTCATGTGTGTACTATTTATCTATTTATTAACGTTTT 731
Db 326 GTTTGGGCAATTTCAATTTCTCATGTGTGTACTATTTATCTATTTATTAACGTTTT 267
Qy 732 AAACGAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGAAATAGCCAGCGGA 791
Db 266 AAACGAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGAAATAGCCAGCGGA 207
Qy 792 TCTCCAGACCAATCTCTCCATGTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 851
Db 206 TTTCCAGACCAATCTCTCCATGTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 147
Qy 852 GCTATAGTGAACATCTCTGCGGCTTCTAGGCTTGTCTCTTATGTTCTTAATCA 911
Db 146 GNTATAGTGAACATCTCTGCGGCTTCTAGGCTTGTCTCTTATGTTCTTAATCA 87
Qy 912 GATACTGCTGGAAGCTTTCAATTTTACAGCCCTGAGAGAGCTTTCTTGTAGTTGA 971
Db 86 GATACTGCTGGAAGCTTTCAATTTTACAGCCCTGAGAGAGTNTTCTTGTAGTTGA 27
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1

```

```

RESULT 6
US-09-352-616A-72/c
; Sequence 72, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:

```

```

; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yugu
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(511)
; OTHER INFORMATION: n = A,T,C or G
US-09-352-616A-72

```

```

Query Match          43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1.7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

Qy 492 CCTGCTCTCAAGAGTGTGTCTATTTTGGCCATCATCCTGCTGTCTTGAAGAAC 551
Db 506 CNGGCTCTCAAGAGGAGTGTATTTTNGCCATCATCCTGNGTTGTTGAAGAACG 447
Qy 552 GCGTTTGTGTCGATGAGAGAAATCATACAGCGGTGACAGAGGCGCATCTTTT 611
Db 446 GGTITNMGTCGATGAGAGAAATCATACAGCGGCGACAGAGGCGCATCTTTT 387
Qy 612 CCTCATCGTTATTTGCTCCCTAGAAAGCGCTTCTGAGATCTAGTTGGCTTTCTTTG 671
Db 386 CTTATCGTTATTTGCTCCCTAGAAAGCGTNTCTGAGATCTAGTTGGCTTTCTTTG 327
Qy 672 GTTTGGGCAATTTCAATTTCTCATGTGTGTACTATTTATCTATTTATTAACGTTTT 731
Db 326 GTTTGGGCAATTTCAATTTCTCATGTGTGTACTATTTATCTATTTATTAACGTTTT 267
Qy 732 AAACGAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGAAATAGCCAGCGGA 791
Db 266 AAACGAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGAAATAGCCAGCGGA 207
Qy 792 TCTCCAGACCAATCTCTCCATGTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 851
Db 206 TTTCCAGACCAATCTCTCCATGTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 147
Qy 852 GCTATAGTGAACATCTCTGCGGCTTCTAGGCTTGTCTCTTATGTTCTTAATCA 911
Db 146 GNTATAGTGAACATCTCTGCGGCTTCTAGGCTTGTCTCTTATGTTCTTAATCA 87
Qy 912 GATACTGCTGGAAGCTTTCAATTTTACAGCCCTGAGAGAGCTTTCTTGTAGTTGA 971
Db 86 GATACTGCTGGAAGCTTTCAATTTTACAGCCCTGAGAGAGTNTTCTTGTAGTTGA 27
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1

```

```

RESULT 7
US-09-232-149A-72/c
; Sequence 72, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE

```

```

RESULT 8
US-09-159-812-72/c
; Sequence 72, Application US/09159812A
; Patent No. 6613872
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.42805
; CURRENT APPLICATION NUMBER: US/09/159,812A
; CURRENT FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: Fastseq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 511
; TYPE: DNA
;

```

```

US-SERIAL 9
US-09-636-215-72/c
: Sequence 72, Application US/09636215
: Patent No. 6620922
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Devin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yugu
: APPLICANT: Henderson, Robert A.
: APPLICANT: Kalos, Michael D.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stolk, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Li, Samuel
: APPLICANT: Wang, Aijun
: APPLICANT: Skeiky, Yasir A.W.
: APPLICANT: Hepner, William
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
: FILE REFERENCE: 210121.42717C17

```

;; CURRENT APPLICATION NUMBER: US/09/636,215
;; CURRENT FILING DATE: 2000-08-10
;; NUMBER OF SEQ ID NOS: 852
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 72
;; LENGTH: 511
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(511)
;; OTHER INFORMATION: n = A,T,C or G
US-09-636-215-72

Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1,7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CTTGCTCTCAAGTGTGTCATATTTTGGCCATCATCAGCTGTCTGCTTGAAGAAC 551
Db 506 CNGGCTCTCAAGAGGGGGGTATTTTNGCCATCATCAGCTGTCTGCTTGAAGAAC 447
Qy 552 GGCCTTCTGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 611
Db 446 GATTNTNGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 387
Qy 612 CCTCATGCTTATTTGTCCTTGAAGAGCTCTTGAAGATCTAGTTGGGCTTTCTTCTGG 671
Db 386 CTTATGCTTATTTGTCCTTGAAGAGCTTCTAGATCTAGTTGGGCTTTCTTCTGG 327
Qy 672 GTTTGGGCAATTCAGTTCTCATGTGTACTATTTATCATATTTGTAATACGTTTTC 731
Db 326 GTTTGGGCAATTCAGTTCTCATGTGTACTATTTATCATATTTGTAATACGTTTTC 267
Qy 732 AAACCAAGGGGCAACAGAGAACTCACTGTATTAATACATGAGAAATGAGCAAGGCGA 791
Db 266 AAACCAAGGGGCAACAGAGAACTCACTGTATTAATACATGAGAAATGAGCAAGGCGA 207
Qy 792 TCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACCAATAGCGCT 851
Db 206 TTTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACCAATAGCGCT 147
Qy 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCTTGTCTCTTAAATGCTTAAATCA 911
Db 146 GNTATAGTGAACATCTCTGCGGCTTCTAGCTTGTCTCTTAAATGCTTAAATCA 87
Qy 912 GATAACGCTGGAAGCTTTCATTTTACAGCCCTGAGAGAGCTCTTGTGCTAGTTGA 971
Db 86 GATAACGCTGGAAGCTTTCATTTTACAGCCCTGAGAGAGCTTGTGCTAGTTGA 27
Qy 972 ATTATGAGTGTGTTTTCGTAATA 997
Db 26 ATTATGAGTGTGTTTTCGTAATA 1

RESULT 10
US-09-685-166A-72/c
; Sequence 72, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Reiter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick

;; APPLICANT: Li, Samuel
;; APPLICANT: Wang, Aifun
;; APPLICANT: Skeiky, Yasir A.W.
;; APPLICANT: Hepler, William
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
;; FILE REFERENCE: 210121.427C21
;; CURRENT APPLICATION NUMBER: US/09/685,166A
;; CURRENT FILING DATE: 2000-10-10
;; NUMBER OF SEQ ID NOS: 898
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 72
;; LENGTH: 511
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(511)
;; OTHER INFORMATION: n = A,T,C or G
US-09-685-166A-72

Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1,7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CTTGCTCTCAAGTGTGTCATATTTTGGCCATCATCAGCTGTCTGCTTGAAGAAC 551
Db 506 CNGGCTCTCAAGAGGGGGGTATTTTNGCCATCATCAGCTGTCTGCTTGAAGAAC 447
Qy 552 GGCCTTCTGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 611
Db 446 GATTNTNGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 387
Qy 612 CCTCATGCTTATTTGTCCTTGAAGAGCTCTTGAAGATCTAGTTGGGCTTTCTTCTGG 671
Db 386 CTTATGCTTATTTGTCCTTGAAGAGCTTCTAGATCTAGTTGGGCTTTCTTCTGG 327
Qy 672 GTTTGGGCAATTCAGTTCTCATGTGTACTATTTATCATATTTGTAATACGTTTTC 731
Db 326 GTTTGGGCAATTCAGTTCTCATGTGTACTATTTATCATATTTGTAATACGTTTTC 267
Qy 732 AAACCAAGGGGCAACAGAGAACTCACTGTATTAATACATGAGAAATGAGCAAGGCGA 791
Db 266 AAACCAAGGGGCAACAGAGAACTCACTGTATTAATACATGAGAAATGAGCAAGGCGA 207
Qy 792 TCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACCAATAGCGCT 851
Db 206 TTTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACCAATAGCGCT 147
Qy 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCTTGTCTCTTAAATGCTTAAATCA 911
Db 146 GNTATAGTGAACATCTCTGCGGCTTCTAGCTTGTCTCTTAAATGCTTAAATCA 87
Qy 912 GATAACGCTGGAAGCTTTCATTTTACAGCCCTGAGAGAGCTCTTGTGCTAGTTGA 971
Db 86 GATAACGCTGGAAGCTTTCATTTTACAGCCCTGAGAGAGCTTGTGCTAGTTGA 27
Qy 972 ATTATGAGTGTGTTTTCGTAATA 997
Db 26 ATTATGAGTGTGTTTTCGTAATA 1

RESULT 11
US-09-115-453-72/c
; Sequence 72, Application US/09115453B
; Patent No. 6657056
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Reiter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick

;; CURRENT FILING DATE: 1998-07-14
;; NUMBER OF SEQ ID NOS: 228
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 72
;; LENGTH: 511
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(511)
;; OTHER INFORMATION: n = A,T,C or G
US-09-115-453-72

Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1,7e-127; Indels 0; Gaps 0;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CCTGCTCTCAAGAGTGTCTATTTTGGCATCATCCTGTCTGCTTAGAAGAAC 551
DB 506 CNGGCTCTCAAGAGGAGGTATTTTNGCATCATCCTGTGCTTAGAAGAAC 447
QY 552 GCGTTTCTGCTCAATGAGAAATCATACAGAGGTGGCACAAGAGGCGCATCTTT 611
DB 446 GGTATTNGCTCAATGAGAAATCATACAGAGGAGGCGCACAAGAGGCGCATCTTT 387
QY 612 CCTCATCGTTATTTGCTCCCTAGAGGCTTTCTGAGATCTAGTTGGCTTTCTTTCTGG 671
DB 386 CTTCATCGTTATTTGCTCCCTAGAGGCTTTCTGAGATCTAGTTGGCTTTCTTTCTGG 327
QY 672 GTTTGGCCATTTCAGTTCTCATGTGTACTATTTATCATATTATTTATTAACGGTTTC 731
DB 326 GTTTGGCCATTTCAGTTCTCATGTGTACTATTTATCATATTATTTATTAACGGTTTC 267
QY 732 AAACAGTGGGACACAGAGAACTCTACTCTGTAATACATAGGAATAGCCAGCGCA 791
DB 266 AAACAGTGGGACACAGAGAACTCTACTCTGTAATACATAGGAATAGCCAGCGCA 207
QY 792 TCTCCAGACCAATCTCTCAGTTTTCACAGCTCTCCAGCCCAACCAATAGCGCT 851
DB 206 TCTCCAGACCAATCTCTCAGTTTTCACAGCTCTCCAGCCCAACCAATAGCGCT 147
QY 852 GCTAATGTAAGACATCTCGGCTTCTAGCCTTCTCTCTCTAGTTCTTTATCA 911
DB 146 GNTATGTAAGACATCTCGGCTTCTAGCCTTCTCTCTCTAGTTCTTTATCA 87
QY 912 GATPACTGCTGGAAGCTTTTATTTTACAGCCCTGAGAGAGTCTTTCTGAGTTGA 971
DB 86 GATPACTGCTGGAAGCTTTTATTTTACAGCCCTGAGAGAGTCTTTCTGAGTTGA 27
QY 972 ATTATGCTGTGTCTTTTCCGTATA 997
DB 26 ATTATGCTGTGTCTTTTCCGTATA 1

RESULT 12
US-09-688-489-72/c
;; Sequence 72, Application US/09688489
;; Patent No. 6664377
;; GENERAL INFORMATION:
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Dillon, Davin C.
;; APPLICANT: Mitcham, Jennifer Lynn
;; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
;; FILE REFERENCE: 210121.427D2
;; CURRENT APPLICATION NUMBER: US/09/488,489
;; CURRENT FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 338
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 72
;; LENGTH: 511
;; TYPE: DNA
;; ORGANISM: Homo sapien

;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(511)
;; OTHER INFORMATION: n = A,T,C or G
US-09-688-489-72

Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1,7e-127; Indels 0; Gaps 0;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CCTGCTCTCAAGAGTGTCTATTTTGGCATCATCCTGTCTGCTTAGAAGAAC 551
DB 506 CNGGCTCTCAAGAGGAGGTATTTTNGCATCATCCTGTGCTTAGAAGAAC 447
QY 552 GCGTTTCTGCTCAATGAGAAATCATACAGAGGTGGCACAAGAGGCGCATCTTT 611
DB 446 GGTATTNGCTCAATGAGAAATCATACAGAGGAGGCGCACAAGAGGCGCATCTTT 387
QY 612 CCTCATCGTTATTTGCTCCCTAGAGGCTTTCTGAGATCTAGTTGGCTTTCTTTCTGG 671
DB 386 CTTCATCGTTATTTGCTCCCTAGAGGCTTTCTGAGATCTAGTTGGCTTTCTTTCTGG 327
QY 672 GTTTGGCCATTTCAGTTCTCATGTGTACTATTTATCATATTATTTATTAACGGTTTC 731
DB 326 GTTTGGCCATTTCAGTTCTCATGTGTACTATTTATCATATTATTTATTAACGGTTTC 267
QY 732 AAACAGTGGGACACAGAGAACTCTACTCTGTAATACATAGGAATAGCCAGCGCA 791
DB 266 AAACAGTGGGACACAGAGAACTCTACTCTGTAATACATAGGAATAGCCAGCGCA 207
QY 792 TCTCCAGACCAATCTCTCAGTTTTCACAGCTCTCCAGCCCAACCAATAGCGCT 851
DB 206 TCTCCAGACCAATCTCTCAGTTTTCACAGCTCTCCAGCCCAACCAATAGCGCT 147
QY 852 GCTAATGTAAGACATCTCGGCTTCTAGCCTTCTCTCTCTAGTTCTTTATCA 911
DB 146 GNTATGTAAGACATCTCGGCTTCTAGCCTTCTCTCTCTAGTTCTTTATCA 87
QY 912 GATPACTGCTGGAAGCTTTTATTTTACAGCCCTGAGAGAGTCTTTCTGAGTTGA 971
DB 86 GATPACTGCTGGAAGCTTTTATTTTACAGCCCTGAGAGAGTCTTTCTGAGTTGA 27
QY 972 ATTATGCTGTGTCTTTTCCGTATA 997
DB 26 ATTATGCTGTGTCTTTTCCGTATA 1

RESULT 13
US-09-679-426-72/c
;; Sequence 72, Application US/09679426
;; Patent No. 6759515
;; GENERAL INFORMATION:
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Dillon, Davin C.
;; APPLICANT: Mitcham, Jennifer L.
;; APPLICANT: Harlocker, Susan L.
;; APPLICANT: Jiang, Yugu
;; APPLICANT: Henderson, Robert A.
;; APPLICANT: Kalos, Michael D.
;; APPLICANT: Fanger, Gary R.
;; APPLICANT: Retter, Marc W.
;; APPLICANT: Stolk, John A.
;; APPLICANT: Day, Craig H.
;; APPLICANT: Vedavick, Thomas S.
;; APPLICANT: Carter, Darrick
;; APPLICANT: Li, Samuel
;; APPLICANT: Wang, Aijun
;; APPLICANT: Skeiky, Yasir A.W.
;; APPLICANT: Hepler, William
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; FILE REFERENCE: 210121.427C20
;; CURRENT APPLICATION NUMBER: US/09/679,426

CURRENT FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 895
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 72
LENGTH: 511
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(511)
OTHER INFORMATION: n = A,T,C or G
US-09-679-426-72

Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1.7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CCTGCTCTCAAGAGTGTCTATTTTGGCATCATCACTGCTGTCTGCTTAGAAGAC 551
Db CNGGTCCTCAAGAGGGGTGTATTTTNGCATCATCACTGNGTTTCTTAGAAGACG 447
Qy 552 GCGTTTCTGCTGCAATGAGAGAAATCATACAGCGGTGCAAGAGGCAATCTTTT 611
Db GGTATTNGCTGCAATGAGAGAAATCATACAGCGGTGCAAGAGGCAATCTTTT 387
Qy 612 CCTCATCGTTATTTGTCCTAGAAAGCGCTTCTGAGATCTAGTTGGCTTTCTTCTG 671
Db CTTATCGGTTATTTGTCCTAGAAAGCGTNTCTGAGATCTAGTTGGCTTTCTTCTG 327
Qy 672 GTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATTCATTATTTGTAACGGTTTC 731
Db GTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATTCATTATTTGTAACGGTTTC 267
Qy 732 AAACAGTGGGCAACAGAGAACTCTGTAATACATAGAGAAATAGCCAGGCA 791
Db AAACAGTGGGCAACAGAGAACTCTGTAATACATAGAGAAATAGCCAGGCA 207
Qy 792 TCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 851
Db TTTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 147
Qy 852 GCTATAGTGAACATCTCGGGCTTCTAGCGCTTGTCCCTCTCTTAGTCTTAAATCA 911
Db GNTATAGTGAACATCTCGGGCTTCTAGCGCTTGTCCCTCTCTTAGTCTTAAATCA 87
Qy 912 GATACTGCTGGAAGCCTTTCATTTTACAGCCCTGGAAGAGCTTCTTTGCTAGTTGA 971
Db GATACTGCTGGAAGCCTTTCATTTTACAGCCCTGGAAGAGTCTTTGCTAGTTGA 27
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997
Db ATTATGTGTGTGTTTTCCGTAATA 1

RESULT 14
US-09-759-143-72/c
Sequence 72, Application US/09759143
Patent No. 6800746
GENERAL INFORMATION:

APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jjiang, Yuqui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 72
LENGTH: 511
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(511)
OTHER INFORMATION: n = A,T,C or G
US-09-759-143-72

Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1.7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CCTGCTCTCAAGAGTGTCTATTTTGGCATCATCACTGCTGTCTGCTTAGAAGAC 551
Db CNGGTCCTCAAGAGGGGTGTATTTTNGCATCATCACTGNGTTTCTTAGAAGACG 447
Qy 552 GCGTTTCTGCTGCAATGAGAGAAATCATACAGCGGTGCAAGAGGCAATCTTTT 611
Db GGTATTNGCTGCAATGAGAGAAATCATACAGCGGTGCAAGAGGCAATCTTTT 387
Qy 612 CCTCATCGTTATTTGTCCTAGAAAGCGCTTCTGAGATCTAGTTGGCTTTCTTCTG 671
Db CTTATCGGTTATTTGTCCTAGAAAGCGTNTCTGAGATCTAGTTGGCTTTCTTCTG 327
Qy 672 GTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATTCATTATTTGTAACGGTTTC 731
Db GTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATTCATTATTTGTAACGGTTTC 267
Qy 732 AAACAGTGGGCAACAGAGAACTCTGTAATACATAGAGAAATAGCCAGGCA 791
Db AAACAGTGGGCAACAGAGAACTCTGTAATACATAGAGAAATAGCCAGGCA 207
Qy 792 TCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 851
Db TTTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 147
Qy 852 GCTATAGTGAACATCTCGGGCTTCTAGCGCTTGTCCCTCTCTTAGTCTTAAATCA 911
Db GNTATAGTGAACATCTCGGGCTTCTAGCGCTTGTCCCTCTCTTAGTCTTAAATCA 87
Qy 912 GATACTGCTGGAAGCCTTTCATTTTACAGCCCTGGAAGAGCTTCTTTGCTAGTTGA 971
Db GATACTGCTGGAAGCCTTTCATTTTACAGCCCTGGAAGAGTCTTTGCTAGTTGA 27
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997
Db ATTATGTGTGTGTTTTCCGTAATA 1

RESULT 15
US-09-651-236-72/c
Sequence 72, Application US/09651236
Patent No. 6818751
GENERAL INFORMATION:

APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jjiang, Yuqui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.


```
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42718C18
; CURRENT APPLICATION NUMBER: US/09/651.236
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 865
; SOFTWARE: FaalSEQ for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(511)
; OTHER INFORMATION: n = A,T,C or G
US-09-651-236-72
```

```
Query Match 43.2%; Score 444; DB 3; Length 511;
Best Local Similarity 91.3%; Pred. No. 1.7e-127;
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
```

```
QY 492 CCGCTCTCTCAAGAGTGTGCTATTTTGGCATCATCACTGCTGCTGCTTGAAGAC 551
Db 506 CNGGTCCTTCAAGAGGGGGGTATTTTNGCATCATCACTGCTGCTTGAAGACG 447
QY 552 GCGTTTCTGCTCAATGAGAGAAATCATTAACAGCGTGGGACAGAGAGCCATCTTTT 611
Db 446 GCTTTTNGCTGCAATGAGAGAAATCATTAACAGCGGGGACAGAGAGCCATCTTTT 387
QY 612 CCTCATGGTTATTTGCTCCCTTGAAGCGCTTCTGAGATCTAGTTGGGCTTTCTTCTGG 671
Db 386 CTTTCATCGTTATTTGCTCCCTTGAAGCGTNTTCTGAGATCTAGTTGGGCTTTCTTCTGG 327
QY 672 GTTGGGCAATTTAGTTCATGTGTACTATTTATCATTTATTTGATTAAGGTTTTC 731
Db 326 GTTGGGCAATTTAGTTCATGTGTACTATTTATCATTTATTTGATTAAGGTTTTC 267
QY 732 AAACGAGTGGGACACAGAGAACTCACTGTATATAAATGAGATAGCCAGGCGA 791
Db 266 AAACGAGTGGGACACAGAGAACTCACTGTATATAAATGAGATAGCCAGGCGA 207
QY 792 TCTCAGACACCAATCTCTCCATGTTTTCCACAGGCTCCTCCAGCCCAATATAGGCGCT 851
Db 206 TCTCAGACACCAATCTCTCCATGTTTTCCACAGGCTCCTCCAGCCCAATATAGGCGCT 147
QY 852 GCTTATAGTGAATCTCTGGGCTTCTAGGCTTGTCCCTCTTATAGTGTCTTAATCA 911
Db 146 GNTATAGTGAATCTCTGGGCTTCTAGGCTTGTCCCTCTTATAGTGTCTTAATCA 87
QY 912 GATAAGTGGTGAAGGCTTCTATTTTACAGCCCTGAAGAGAGCTTCTTGTAGTTGA 971
Db 86 GATAAGTGGTGAAGGCTTCTATTTTACAGCCCTGAAGAGAGTCTTGTAGTTGA 27
QY 972 ATTATGTGTGTGTTTTCCGTAATA 997
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1
```

Search completed: December 10, 2005, 23:27:03
Job time : 232 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 20:04:37 ; Search time 4078 Seconds
(without alignments)
11782.818 million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027
Sequence: 1 gggcgaagctggcgcaaaa.....atttaaaatgaagaact 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : EST.*
Listing first 45 summaries

1: gb_esc1.*
2: gb_esc2.*
3: gb_esc3.*
4: gb_hnc.*
5: gb_esc4.*
6: gb_esc5.*
7: gb_esc6.*
8: gb_esc7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	# Query	Match Length	DB ID	Description
1	961.8	93.7	3533	4	BC030554 Homo sapi
2	880.8	85.8	1510	4	BC017861 Homo sapi
3	837.8	81.6	914	5	BO934484 AGNCOURT
4	636	61.9	821	2	BF679123 602153390
5	627.8	61.1	997	2	BF678971 602153608
6	618.6	60.2	636	1	BF674593 602137811
7	614	59.8	938	1	BF674593 602137811
8	613.8	59.8	938	1	BF674593 602137811
9	605.2	58.9	959	2	CN645427 ILUMIGEN
10	590.2	57.5	703	2	BC217853 RSM73574
11	589.6	57.4	629	1	AT168834 wj03h04.x
12	584	56.9	700	6	CA449324 UI-H-E11-
13	580.4	56.5	582	2	BS326754 h164d05.x
14	578.8	56.4	613	3	BI905966 603062856
15	569.8	55.5	719	2	BF681238 602155502
16	567.8	55.3	670	2	BF677915 602084736
17	565.6	55.1	789	5	BU199143 DCBCIA08
18	561	54.6	629	2	BF679165 602153737
19	558.4	54.4	1058	3	BM544213 AGNCOURT
20	553	53.8	849	2	BF677648 602085529
21	552.6	53.8	849	2	BF679284 602153477
22	548.8	53.4	803	2	BF680840 602156088

c	23	547.6	53.3	701	10	AG175832 Pan tlog1
	24	543.6	52.9	857	2	BF674457 BF674457 602137231
	25	540.2	52.6	861	2	BF681385 BF681385 602156579
c	26	532.4	51.8	665	1	AA569813 AA569813 nm46c02.8
c	27	530.8	51.7	600	1	AT244217 AT244217 q186c05.x
	28	516	50.2	1055	2	BF694904 BF694904 602268596
	29	493.4	48.0	927	2	BF679476 BF679476 602153280
	30	490.8	47.8	572	2	BF370026 BF370026 MR3-FN00
c	31	490.4	47.8	503	1	AT253507 AT253507 ag33c11.x
	32	490	47.7	525	5	EX112931 EX112931
	33	488.4	47.6	513	1	AI823897 wj28e10.x
c	34	466	45.4	479	1	AM575113 UI-HF-BKO
c	35	449.4	43.8	596	1	AV715641 AV715641
c	36	438.6	43.7	529	1	AA908306 AA908306 qg33h07.8
c	37	435.6	42.4	442	1	AI718234 AI718234 ag42h05.x
c	38	434.8	42.3	497	1	AA654656 AA654656 nc76c02.8
c	39	429.2	41.8	512	1	AA657507 AA657507 nc64g07.8
c	40	428.6	41.7	681	1	AV714756 AV714756
	41	428.4	41.7	870	2	BF678129 BF678129 602085181
c	42	428	41.7	460	2	BE772783 BE772783 RCI-PT013
c	43	420.4	40.9	965	1	AI557112 AI557112 PT2.1.13
	44	418.6	40.8	523	2	BF675605 BF675605 602083495
c	45	416.8	40.6	457	8	N50880 YY92d06.s1

ALIGNMENTS

RESULT 1
BC030554
LOCUS
DEFINITION
Homo sapiens, similar to T cell receptor gamma locus, clone IMAGE:5213435, mRNA.
BC030554.1 GI:20988582
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 3533)
Strausberg, R.
DIRECT SUBMISSION
Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
REMARK
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Bakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Leric, P., Legaepi, R., Maduro, O.L., Masillo, C., Masker, B., Maestrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Tsungson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 64 Row: k Column: 18
This clone has the following problem: frame shifted.
Location/Qualifiers

FEATURES

	source	1..3533	/organism="Homo sapiens"
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5213435"	
		/tissue_type="Blood, adult leukocytes"	
		/clone_id="NH MG_C_118"	
		/lab_host="DH10B"	
		/note="Vector: pCMV-SPORT6"	
ORIGIN			
Query Match	93.7%	Score 961.8;	DB 4; Length 3533;
Best Local Similarity	99.4%	Pred. No. 7.9e-252;	
Matches	963;	Conservative 0;	Mismatches 6; Indels 0; Gaps 0;
Dy	59	CAGATAAACAACCTGATGCAGATGTTTCCCCAAGCCCACTATTTCCTCCTCATTG	118
Dd	743	CTGTATAAACCACTGGATGCAGATGTTTCCCCCAAGCCCACTATTTCCTCCTCATTG	802
Dy	119	CTGAACAACAACCTCGAGAAGCTGGAATATACCTTGTCTTGGAGAAATTTTTCCCTG	178
Dd	803	CTGAACAACAACCTCGAGAAGCTGGAATATACCTTGTCTTGGAGAAATTTTTCCCTG	862
Dy	179	ATGTTATTAAGATACATTGGCAAAGAAAAGAGCAACGATTCTGGAGATCCCGAGAG	238
Dd	863	ATGTTATTAAGATACATTGGCAAAGAAAAGAGCAACGATTCTGGAGATCCCGAGAG	922
Dy	239	GGAACACCATGAAGACTAACGACACATACATGAATTTTAGCTGTTAACGTTGCCAGAA	298
Dd	923	GGAACACCATGAAGACTAACGACACATACATGAATTTTAGCTGTTAACGTTGCCAGAA	982
Dy	299	AGTCACTGGAACAAAGAACACAGATGTATTCGTACACATGAGAAATAATAAAAAGGATG	358
Dd	983	AGTCACTGGAACAAAGAACACAGATGTATTCGTACACATGAGAAATAATAAAAAGGATG	1042
Dy	359	ATCAAGAAATATATCTTCTCTCCAATPAAGAAGAGTATCATCAATGGATCCCAAGACA	418
Dd	1043	ATCAAGAAATATATCTTCTCTCCAATPAAGAAGAGTATCATCAATGGATCCCAAGACA	1102
Dy	419	ATTGTTCAAAAGATGCAATATGATACATCACTGCTGCACTCAAAACCTCTGCATATT	478
Dd	1103	ATTGTTCAAAAGATGCAATATGATACATCACTGCTGCACTCAAAACCTCTGCATATT	1162
Dy	479	ACATGATCCTCTCTCTGCTCCTCAAGAGTGTGTATTTTGGCATCATCACCTGCTGTC	538
Dd	1163	ACATGATCCTCTCTCTGCTCCTCAAGAGTGTGTATTTTGGCATCATCACCTGCTGTC	1222
Dy	539	TGCTTAGAAGAACGGCTTTCGTGCTGCAATGGAGAAATCATPAACAGCGGTGGCAAG	598
Dd	1223	TGCTTAGAAGAACGGCTTTCGTGCTGCAATGGAGAAATCATPAACAGCGGTGGCAAG	1282
Dy	599	GAGGCAATCTTTCCTCATCGGTTATGTTCCTTAGAAGCGTTCCTGAGATCTAGTTGG	658
Dd	1283	GAGGCAATCTTTCCTCATCGGTTATGTTCCTTAGAAGCGTTCCTGAGATCTAGTTGG	1342
Dy	659	GCTTTCCTTTCGGGTTTGGGCAATTCAGTTCTCATGTGTGTACTATTCTATCATTTTG	718
Dd	1343	GCTTTCCTTTCGGGTTTGGGCAATTCAGTTCTCATGTGTGTACTATTCTATCATTTTG	1402
Dy	719	TATAACGGTTTCAAAACCAAGTGGGCAACAGAAACCTCAGCTCTGTATAACATAGAGA	778
Dd	1403	TATAACGGTTTCAAAACCAAGTGGGCAACAGAAACCTCAGCTCTGTATAACATAGAGA	1462
Dy	779	ATAGCGACGGGAGATCCAGACCAATCTCTCAGATGTTTCCAGAGCTCTCCAGCCAAAC	838
Dd	1463	ATAGCGACGGGAGATCCAGACCAATCTCTCAGATGTTTCCAGAGCTCTCCAGCCAAAC	1522
Dy	839	CCAAATAGCGCTCTGATATAGTGTAGACATCTTGCGGCTTCTAGCTGTCTCTCTTAG	898
Dd	1523	CCAAATAGCGCTCTGATATAGTGTAGACATCTTGCGGCTTCTAGCTGTCTCTCTTAG	1582
Dy	899	TGTTCTTTAATCAGATTAACCTGCTGGAAGCTTTCAATTTTACAGCGCTGGAAGAGTCTT	958

Db	1583	TGTTCTTTAATCAGATACGCGCGAAGCCTTTCATTTCACGCGCCTCAGACAGCTT	1642
Qy	959	CTTTGCTGACTGGAATTTATGTCGTGTGTTTTTCGTAATTAAGCAAAATTAATTTAAAAAA	1018
Db	1643	CTTTCGAGTTGGAATTAATGTCGTGTGTTTTTCGTAATTAAGCAAAATTAATTTAAAAAA	1702
Qy	1019	TGAAAAGTT 1027	
Db	1703	TGAAAAGTT 1711	
RESULT 2			
LOCUS	BC017861	1510 bp	mRNA
DEFINITION	Homo sapiens, similar to T cell receptor gamma locus, clone		
ACCESSION	BC017861		
VERSION	BC017861		
KEYWORDS	HTC.		
SOURCE	BC017861.1 GI:17389678		
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
	1 (bases 1 to 1510)		
REFERENCE	Strausberg, R.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-DEC-2001) National Institutes of Health, Mammalian		
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
REMARK	Contact: MGC help desk		
COMMENT	Email: cgabs-remail.nih.gov		
	Tissue Procurement: CLONTECH		
	CDNA Library Preparation: CLONTECH Laboratories, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome		
	Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: http://www-shgc.stanford.edu		
	Contact: (Dickson, Mark) mcd@paxil.stanford.edu		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,		
	R. M.		
	Clone distribution: MGC clone distribution information can be found		
	through the I.M.A.G.E. Consortium/LILN at: http://image.liln.gov		
	Series: IRAL Plate: 36 Row: k Column: 19		
	This clone has the following problem: frame shifted.		
FEATURES	Location/Qualifiers		
Source	1..1510		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4687960"		
	/tissue_type="Lung"		
	/clone_id="NIH_MGC_77"		
	/lab_host="DH10B"		
	/note="Vector: pDNR-LIB"		
ORIGIN			
	Query Match 85.8%; Score 880.8; DB 4; Length 1510;		
	Best Local Similarity 93.6%; Pred. No. 1e-229;		
	Matches 949; Conservative 0; Mismatches 17; Indels 48; Gaps 1;		
Qy	59	CAGATTAACAAGCTGATGACAGATGTTTCCCAAGCCACATATTTCTTCTCAATG 118	
Db	464	CTGATTAACAAGCTGATGACAGATGTTTCCCAAGCCACATATTTCTTCTCGATTG 523	
Qy	119	CTGAAAACAAAGCTCAGAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTG 178	
Db	524	CTGAAAACAAAGCTCAGAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCAG 583	
Qy	179	ATGTTATTAGATACATTGCGAAGAAAAGAGCAACAGATTTCTGGATCCACAGAGG 238	

```

Db      584 ATATTATTAAGATATGATGGCAAGAAAGAGCAACAGATCTCGATCCAGAGG 643
Qy      239 GGAACACCATGAAAGACTTAACGACATACATGAAATTTAGCTGTTAACGGTCCAGAA 298
Db      644 GGAACACCATGAAAGACTTAACGACATACATGAAATTTAGCTGTTAACGGTCCAGAA 703
Qy      299 AGTACATGGAACAAGAACACAGATGTATGTCAGACATGAGAAATTAATTAAGACGAGTTG 358
Db      704 AGTACATGGAACAAGAACACAGATGTATGTCAGACATGAGAAATTAATTAAGACGAGTTG 763
Qy      359 ATCAAGAAATTAATCTTCTCTCAATTAAGAC-----389
Db      764 ATCAAGAAATTAATCTTCTCTCAATTAAGACATGTCACACAGTGGATCCAAATACA 823
Qy      390 -----GGATGTCAATCAATGATGATCCCAAGACAAATTTTCAAAAG 430
Db      824 ATTAATCAAGATGCAAAATGATGTCAATGATCCCAAGACAAATTTGTCAAAG 883
Qy      431 ATGAATGATGATACATGCTGCTGAGCTCAAAACACCTGCTGATTAATGATGATCTCC 490
Db      884 ATGAATGATGATACATGCTGCTGAGCTCAAAACACCTGCTGATTAATGATGATCTCC 943
Qy      491 TCCTGCTCTCAAGAGTGTGTCTATTTGTCATCATCACTGCTGCTGCTGCTGCTGCTGCTG 550
Db      944 TCCTGCTCTCAAGAGTGTGTCTATTTGTCATCATCACTGCTGCTGCTGCTGCTGCTGCTG 1003
Qy      551 CGGCTTCTGCTGCAATGAGAGAAATCATPAACAGACGGTGGCAAGAGAGGCACTTTT 610
Db      1004 CGGCTTCTGCTGCAATGAGAGAAATCATPAACAGACGGTGGCAAGAGAGGCACTTTT 1063
Qy      611 TCCTCATCGGTATTTGTCCTGAGAAAGGCTCTGAGAGATCTAGTGGCTTTCTTTCTG 670
Db      1064 TCCTCATCGGTATTTGTCCTGAGAAAGGCTCTGAGAGATCTAGTGGCTTTCTTTCTG 1123
Qy      671 GGTGGGCAATTCAGTCTCATGTGTGATCTATTCATCATTTATGTAATGCGTTT 730
Db      1124 GGTGGGCAATTCAGTCTCATGTGTGATCTATTCATCATTTATGTAATGCGTTT 1183
Qy      731 CAACACAGTGGGCAACAGAGAACTCTGTATTAATPAACAGAGAAATGACAGCGG 790
Db      1184 CAACACAGTGGGCAACAGAGAACTCTGTATTAATPAACAGAGAAATGACAGCGG 1243
Qy      791 ATCTCCAGACCAATCTCTCATGTGTTTCCAGACTCTCTCAGACCAACCCAAATAGGCC 850
Db      1244 ATCTCCAGACCAATCTCTCATGTGTTTCCAGACTCTCTCAGACCAACCCAAATAGGCC 1303
Qy      851 TGCTATGTGATGACATCTGCGGCTTCTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910
Db      1304 TGCTATGTGATGACATCTGCGGCTTCTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1363
Qy      911 AGATTAAGTCTGGAAGGCTTTCAATTTTACAGCCCTGAGACAGCTCTTGTGCTAGTGG 970
Db      1364 AGATTAAGTCTGGAAGGCTTTCAATTTTACAGCCCTGAGACAGCTCTTGTGCTAGTGG 1423
Qy      971 AATTATGTGTGTGTTTTCGTAATAAGCAAAATTAATTTAAAAAATGAAA 1024
Db      1424 AATTATGTGTGTGTTTTCGTAATAAGCAAAATTAATTTAAAAAATGAAA 1477

```

```

REFERENCE 1 (bases 1 to 914)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M13628 row: c column: 11
High quality sequence stop: 633.
FEATURES
Source
1. 914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205378"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski, sciatic nerve"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTCTTGAATCGGAGCGGCCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
ORIGIN
Query Match 81.6%; Score 837.8; DB 5; Length 914;
Best Local Similarity 98.5%; Pred. No. 5.4e-218;
Matches 856; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
Qy 157 CTCTTGAGAAATTTTTCCTGATGTTATTAAGATACATTGCAAGAAAGAGCAAC 216
Db 1 CTCTTGAGAAATTTTTCCTGATGTTATTAAGATACATTGCAAGAAAGAGCAAC 60
Qy 217 ACGATTCGGGATCCAGAGGGGAAACCATGAAGCTAAGCAATACATGAAATTT 276
Db 61 ACGATTCGGGATCCAGAGGGGAAACCATGAAGCTAAGCAATACATGAAATTT 120
Qy 277 AGCTGGTTAAGGTGCCAGAAAGTCACTGAGCAAAAGACAGATGTATCGTCAAGCAT 336
Db 121 AGCTGGTTAAGGTGCCAGAAAGTCACTGAGCAAAAGACAGATGTATCGTCAAGCAT 180
Qy 337 GAGATTAATAAAACGAGGTGATCAAGAAATTTATCTTCTCAATTAAGACGATGTC 396
Db 181 GAGATTAATAAAACGAGGTGATCAAGAAATTTATCTTCTCAATTAAGACGATGTC 240
Qy 397 ATCAATGATCCCAAGACAAATTTGTTCAAAAGATGCAATATGATCACTACTGTCAG 456
Db 241 ATCAATGATCCCAAGACAAATTTGTTCAAAAGATGCAATATGATCACTACTGTCAG 300
Qy 457 CTCACAAACCTCTGATATTACATGATCTCTCTGCTCTCAAGATGTGTTAT 516
Db 301 CTCACAAACCTCTGATATTACATGATCTCTCTGCTCTCAAGATGTGTTAT 360
Qy 517 TTGGCATCATCACTGCTGCTGCTTGAAGAACGCTTCTGCTGCAATGAGAGAAA 576
Db 361 TTGGCATCATCACTGCTGCTGCTTGAAGAACGCTTCTGCTGCAATGAGAGAAA 420
Qy 577 TCATTAACAGACGGTGGCAAGAGAGGCATCTTTCTCATCGGTATTTGCTTAGAG 636
Db 421 TCATTAACAGACGGTGGCAAGAGAGGCATCTTTCTCATCGGTATTTGCTTAGAG 480

```

RESULT 3
 B0934484
 LOCUS
 DEFINITION
 AGNCOURT 8754342 Lupski, sciatic nerve Homo sapiens cDNA clone
 IMAGE:6205378 5', mRNA sequence.
 B0934484
 ACCESSION
 B0934484.1 GI:22349867
 VERSION
 B0934484
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

Qy	637	CGCTCTGAGAGATCTAGTGGGCTTTCTTTGCGGTTTGGCCATTCAATCTTCATGT	696
Db	481	CGCTCTGAGAGATCTAGTGGGCTTTCTTTGCGGTTTGGCCATTCAATCTTCATGT	540
Qy	697	GTGTACTATTCATCATTTATGTATPACGGTTTCAACACAGTGGGACACAGAGAACCT	756
Db	541	GTGTACTATTCATCATTTATGTATPACGGTTTCAACACAGTGGGACACAGAGAACCT	600
Qy	757	CACCTCTGTATATPACAATGAGGAATAGCCACGGCGATCTCCACACCAATCTCTCCATGTT	816
Db	601	CACCTCTGTATATPACAATGAGGAATAGCCACGGCGATCTCCACACCAATCTCTCCATGTT	660
Qy	817	TTCCACAGGTCCTCCAGGCCAACCCAAATPAGGCGCTGTAGTGTAGACATCTGGGCGCT	876
Db	661	TTCCACAGGTCCTCCAGGCCAACCCAAATPAGGCGCTGTAGTGTAGACATCTGGGCGCT	720
Qy	877	TCTAGCCTTGTCCCTCTCTCTAGTGTCTTTAATTCAGATPACTGCTGGAAGCTTTCATT	936
Db	721	TCTAGCCTTGTCCCTCTCTCTAGTGTCTTTAATTCAGATPACTGCTGGAAGCTTTCATT	780
Qy	937	TTACACGCCCTGAGAGAGTCTTCTTTGCTGATTTGAATATGTGT- GTGTTTTCCGTAA	995
Db	781	TTACACGCCCTGAGAAACAGCTTCTTTGCTGATTTGAATATGTGTGGGGGTTTTTCGTAA	840
Qy	996	TAAGCAATAATTAATTAATAATAAATAA 1024	
Db	841	TAAGCAATAATTAATTAATAATAAATAA 869	

```
RESULT 4  
BF679123  
LOCUS  
DEFINITION      BF679123                821 bp    mRNA   linear     EST-21-DEC-2000  
                  60215339001 NIH_MGC_83 Homo sapiens CDNA clone IMAGE:4294247 5',  
RNA sequence.  
ACCESSION       BF679123  
VERSION         BF679123.1 GI:11953018  
KEYWORDS        EST.  
SOURCE          Homo sapiens (human)  
ORGANISM        Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
               1 (bases 1 to 821)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: csapbs-re@mail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LNCMI43 row: d column: 24  
High quality sequence stop: 620.
```

FEATURES

SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294247"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NH MGC 83"
/notes="Organ: prostate; Vector: pDMR-LB (Clontech);
Site_1: S11 (ggcgctcgcgcgc); Site_2: S11
(ggcgcattatggc); 5' and 3' adaptors in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTAGGCC-3'
and 3' adaptor sequence: 5'-ATTCTAGAGCCCGAGCCGCGCAGATG-dt(30)BN-3' (where B = A,
5'-ATTCTAGAGCCCGAGCCGCGCAGATG-dt(30)BN-3' (where B = A,
C, G, or T). Average insert size 1.4

```

ORIGIN

Query Match	61.9%;	Score 636;	DB 2;	Length 821;
Best Local Similarity	95.4%;	Pred. No. 1e-162;		
Matches 731; Conservative	0;	Mismatches 25;	Indels 10;	Gaps 7

kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto CA). "

OY	12	GGGCAAAAATTCAGAGATTTGGTCCCGGAAACAACCTATCACTTCAGATTAACAAC	71
Db	3	GGGCAAAAATTCAGGTA-TTGGTCCCGGAAACAACCTATCACTTCAGATTAACAAC	61
OY	72	TGATGCAGATGTTCCCCCAAGCCCACTATTTTCTTCCTCAATTCGTAAACAAGCT	131
Db	62	TGATGCAGATGTTCCCCCAAGCCCACTATTTTCTTCCTCAATTCGTAAACAAGCT	121
OY	132	CCAGAAGCTGGAACATACCTTTTGCTTTTGAGAAATTTTCCCTGATGTATTAAGT	191
Db	122	CCAGAAGCTGGAACATACCTTTTGCTTTTGAGAAATTTTCCCTGATGTATTAAGT	181
OY	192	ACATTGGCAAGAAAAGAGCAACAGATTCTGGATGCCAGAGGGGAAACACATGAA	251
Db	182	ACATTGGCAAGAAAAGAGCAACAGATTCTGGATGCCAGAGGGGAAACACATGAA	241
OY	252	GACTAACGACATACATGAAATTTAGCTGGTTAACGGTGCAGAAAAAGTCACGTGACAA	311
Db	242	GACTAACGACATACATGAAATTTAGCTGGTTAACGGTGCAGAAAAAGTCACGTGACAA	301
OY	312	AGAACACAGATGTATCGTCAACATGAGAAATATATAAAAACGAGTTGATCAAGAAATAT	371
Db	302	AGAACACAGATGTATCGTCAACATGAGAAATATATAAAAACGAGTTGATCAAGAAATAT	361
OY	372	CTTTCCTCCAAATAAGACGATGTCAATCAACAATGGATCCCAAGACAAATTTGTCAAAA	431
Db	362	CTTTCCTCCAAATAAGACAGATGTCAATCAACAATGGATCCCAAGACAAATTTGTCAAAA	421
OY	432	TGCAAAATGATACATACCTGCTGACGCTCACAACACCTGTGCATATTAACATGATCCTCT	491
Db	422	TGCAAAATGATACATACCTGCTGACGCTCACAACACCTGTGCATATTAACATGATCCTCT	481
OY	492	CTGTGCTCTCAAGAGTGTGTCTATTTTGGCATATCACTGCTGTCTGTATTAAGAAC	551
Db	482	CTGTGCTCTCAAGAGTGTGTCTATTTTGGCATATCACTGCTGTCTGTATTAAGAAC	541
OY	552	GGCTTTTCGTCGACATGAGAG--GAAATCATAAACAGACGGTGGACAA--GGAAGGCATCT	608
Db	542	GGCTTTTCGTCGACATGAGAGAAATTCATTAACAGACGGTGGACAAAGGAGGCATCT	601
OY	609	TTTCTCTACGGTATATGTCCCTAG--AAGCGTTCCTGAGAGCTAGTTGGGCT-TTCT	665
Db	602	TTTCTCTACGGTATATGTCCCTAGATAGGGGCTCTTCTGAAGAGCTAGTTGGGCTATCCT	661
OY	666	TTCTGGGTTTGGGCAATTTCA-GTTCTCATGTGTGTACTATTCTATCATTAATTTGTATAC	724
Db	662	TACTGGGTTAAGGCAATTCACGTGCCCATGTGTGTCTATTCTATCATTAATTTGTATAC	721
OY	725	GGTTTTCACACAGTGGGCAACAGAGAACCTCACTGTATAAC770	
Db	722	GG--TTACACACATTTGGGCAACAGAGAACTCACTGTATATAC765	

RESULT 5	Bf678971	Bf678971	997 bp	mrna	linear	EST 21-DEC-2000
LOCUS	Bf678971	602155608b1	NIH_MGC_83	Homo sapiens	CDNA clone	IMAGE:4294873 5',
DEFINITION		mRNA sequence.				
ACCESSION	Bf678971	Bf678971				
VERSION		Bf678971.1	GI:11952866			
KEYWORDS		EST.	sapiens	(human)		
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 997)
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L10M1144 row: 0 column: 02
High quality sequence stop: 656.
Location/Qualifiers

1. 997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294873"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: Stit (ggcgcctcgcc); Site_2: Stit
(ggcctacggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCGCATATGACC-3',
and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN

Query Match 61.1%; Score 627.8; DB 2; Length 997;
Best Local Similarity 95.0%; Pred. No. 1.9e-160;
Matches 691; Conservative 0; Mismatches 32; Indels 4; Gaps 4;

QY 225 GGGATCCAGAGAGGGAACACATGAGACATACATCATGAAATTTACTGTT 284
DB 1 GGGATCCAGAGAGGGAACACATGAGACATACATCATGAAATTTACTGTT 60
QY 285 AACGTCGCAAGAAAGTCACTGACAAAGAACACAGATGTATCGTCAGACATGAGATTA 344
DB 61 AACGTCGCAAGAAAGTCACTGACAAAGAACACAGATGTATCGTCAGACATGAGATTA 120
QY 345 TAAAAACGAGTTGATCAAGAAATTTATCTTCTCCATATAAGACGATGTCATCAGAT 404
DB 121 TAAAAACGAGTTGATCAAGAAATTTATCTTCTCCATATAAGACGATGTCATCAGAT 180
QY 405 GGATCCCAAGCAATTTGTTCAAAAGATGCAATGATCACTACTGTCGACGCTCACAA 464
DB 181 GGATCCCAAGCAATTTGTTCAAAAGATGCAATGATCACTACTGTCGACGCTCACAA 240
QY 465 CACCTGCAATTAACATGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 524
DB 241 CACCTGCAATTAACATGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 525 CATCACCTGCTGTCTGTAGAGAAACGGCTTCTGTCGAA-TGAGAGAAATCATTAAC 583
DB 301 CATCACCTGCTGTCTGTAGAGAAACGGCTTCTGTCGCAATTTGAGAGAAATCATTAAC 360
QY 584 AGAGGTGGCAAGAGAGGCAATCTTCTCTCACTGATATTTGCTTCCCTAGAGAGGCTCTC 643
DB 361 AGAGGTGGCAAGAGAGGCAATCTTCTCTCACTGATATTTGCTTCCCTAGAGAGGCTCTC 420
QY 644 TGAGGATTAAGTGGGCTTCTCTCTGAGGTTGGGCAATTCAGTTCTCATGTGTAAT 703

DB 421 TGAGGATTAAGTGGGCTTCTCTCTGAGGTTGGGCAATTCAGTTCTCATGTGTAAT 480
QY 704 ATTCTATCATTAATGATTAACGGTTTCAACAGATGGGACACAGAACTCATCTG 763
DB 481 ATTCTATCATTA-TGATTAACGGTTTCAACAGATGGGACACAGAACTCATCTG 539
QY 764 TAATAACATGAGAAATAGCCAGGATCTCCAGACCAATCTCCATGTTTCCACA 823
DB 540 TAATAACATGAGAAATAGCCAGGATCTCCAGACCAATCTCCATGTTTCCACA 598
QY 824 GCTCTCCAGCAACCAATAGCCGCTCTATATGATGACATCTCGGCTTCCAGC 883
DB 599 GCTCTCCAGCAACCAATAGCCGCTCTATATGATGACATCTCGGCTTCCAGC 658
QY 884 TTGTCCTT-CTCTAGTGTCTTATTCAGATATGCTCGGCAAGCTTCAATTTACAC 942
DB 659 TTGTCCTTCTCTTAAGGCTTGTAGCCAGATATGCTGTGGCTTCAATTTACAC 718
QY 943 GCCCTGA 949
DB 719 GCCCTGA 725

RESULT 6
A1972955/c
LOCUS
DEFINITION
A1972955 636 bp mRNA linear EST 08-MAR-2000
w454905.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2490680 3'
similar to gb:U13231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN);,
mRNA sequence.
A1972955
VERSION
A1972955.1 GI:5769781
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 636)
-NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.dio.llnl.gov/dbtp/image/image.html
Insert length: 1305 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 463..
Location/Qualifiers

1. 636
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2490680"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and as
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
985608-986759, 1101192-1101959, and 1217928-1220615).

FEATURES
source

Subtraction by Bento Soares and M. Fatima Bonaldo.

Query Match	60.2%	Score 618.6	DB 1	Length 636
Best Local Similarity	99.2%	Pred. No. 5.7e-158		
Matches 632	Conservative	0	Mismatches 4	Indels 1
				Gaps 1

OY	391	GATGTCATCACAATGATGCCAAMAACAATGTTCTAAAAGATGCAAAATGATCTACCTACG	450
Db	636	GATGTCATCACAATGATGCCAAMAACAATGTTCTAAAAGATGCAAAATGATCTACCTACG	577
OY	451	CTGCAGCTCACAACAACCTCTGCATATTATCATGTACCTCTCTGCTCCTCAAGAGTGTG	510
Db	576	CTGCAGCTCACAAMAAC-CCTCTGCATATTATCAATGTACCTCTCTGCTCCTCAAGAGTGTG	518
OY	511	GTCATTTTGGCATATCATCAACCGCTGTCTGTCTTAAGAAGCGGCTTTCGCTGCAGTAAGGA	570
Db	517	GTCATTTTGGCATATCATCAACCGCTGTCTGTCTTAAAGAAGCGGCTTTCGCTGCAGTAAGGA	458
OY	571	GAGAAATCATTAACAGACGGTGGACAGAAGAGGCATCTTTTCCTCATCGGTATTGTGCC	630
Db	457	GAGAAATCATTAACAGACGGTGGACAGAAGAGGCATCTTTTCCTCATCGGTATTGTGCC	398
OY	631	TAGAAAGCGTCTTCTGAAGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCAATTTCAATTC	690
Db	397	TAGAAAGCGTCTTCTGAAGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCAATTTCAATTC	338
OY	691	TCATGTGTGTAATTTCTATCATTAATTGTAATAAGGTTTCAAAACAGTGGGACACACAGA	750
Db	337	TCATGTGTGTAATTTCTATCATTAATTGTAATAAGGTTTCAAAACAGTGGGACACACAGA	278
OY	751	GAACCTCACTCTGTATAAACAATGAGAAATAGCAACGGGATCTCCAGCAACCAATCTCTC	810
Db	277	GAACCTCACTCTGTATAAACAATGAGAAATAGCAACGGGATCTCCAGCAACCAATCTCTC	218
OY	811	CATGTTTTCCACAGCTCCTCCAGCCCAACCCAAATAGCGCTCTATAGTATGACATCTCT	870
Db	217	CATGTTTTCCACAGCTCCTCCAGCCCAACCCAAATAGCGCTCTATAGTATGACATCTCT	158
OY	871	GCGGCTTCTAGCCTTGTCCTCTCTTAAGTCTTTTAATCAGATAAATCCTCGAAGCCT	930
Db	157	GCGGCTTCTAGCCTTGTCCTCTCTTAAGTCTTTTAATCAGATAAATCCTCGAAGCCT	98
OY	931	TTGATTTTAAACAGCCCTGAAGAGTCTTCTTTGCTAGTGAATATATGTGTGTGTTTTTC	990
Db	97	TTGATTTTAAACAGCCCTGAAGAGTCTTCTTTGCTAGTGAATATATGTGTGTGTTTTTC	38
OY	991	CGTAATTAAGCAAAATTAATTTAAAAAAAAGAAAGTT 1027	
Db	37	CGTAATTAAGCAAAATTAATTTAAAAAAAAGAAAGTT 1	

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMH091 row: a column: 04
High quality sequence step: 629.

ORIGIN

Query Match	59.8%	Score 614	DB 2	Length 902
Best Local Similarity	95.5%	Pred. No. 1.1e-156		
Matches 707; Conservative	0	Mismatches 25	Indels 8	Gaps 7

QY	12	GGGCAAAAATTAACAGGATATTGGTCCCGGAAACAAAGTTATCATTAAGATTAACAT	71
Db	2	GGGCAAAAATTAACAGGATATTGGTCCCGGAAACAAAGTTATCATTAAGATTAACAT	61
QY	72	TGATGCAGATGTTTCCGCCAAGCCACATATTTTCTTCAATTGTGTGAACCAAGCT	131
Db	62	TGATGCAGATGTTTCCGCCAAGCCACATATTTTCTTCAATTGTGTGAACCAAGCT	121
QY	132	CCAGAGGCTGGAACTACCTCTTGTCTTCTTGAGAAATTTTTCCTGATGTTATTAAGT	191
Db	122	CCAGAGGCTGGAACTACCTTGTCTTCTTGAGAAATTTTTCCTGATGTTATTAAGT	181
QY	192	ACATTGGCAAGAAAAGAGACAACGATTCGGGATCCCGAGAGGGGAAACACCATGAA	251
Db	182	ACATTGGCAAGAAAAGAGACAACGATTCGGGATCCCGAGAGGGGAAACACCATGAA	241
QY	252	GACTAACGACACATACATGAAATTTTAACTGTGTTAAAGTGCCAGAAAATCTAGACAA	311
Db	242	GACTAACGACACATACATGAAATTTTAACTGTGTTAAAGTGCCAGAAAATCTAGACAA	301
QY	312	AGAACACAGATGTATCTTCACACATGANAATTAATTAACGAGTTGATCAAGAAATAT	371
Db	302	AGAACACAGATGTATCTTCACACATGANAATTAATTAACGAGTTGATCAAGAAATAT	361
QY	372	CTTTCCTCCATTAAGACGAGTGTTCATACAAATGGATCCCAAGACAA--TTGTTCAAAAG	430
Db	362	CTTTCCTCCATTAAGACAGATGTTCATACAAATGGATCCCAAGACAAATTTGTTCAAAAG	421
QY	431	ATGCMAATGATACACTACTCTGTGCAGCTCAAAACACCTCTGCATATTAATCATGTAACCTCC	490
Db	422	ATGCMAATGATACACTACTCTGTGCAGCTCAAAACACCTCTGCATATTAATCATGTAACCTCC	481
QY	491	TCCGCTCTCTCAAGAGTGGTCTATTTTGCATCATACCTGCTGTCTGTATAAGAA	550
Db	482	TCCGCTCTCTCAAGAGTGGTCTATTTTGCATCATACCTGCTGTCTGTATAAGAA	541
QY	551	CGGCTTTCGTGTCAGATGAGAGAAAT--CATTAACGAGTGGCAAGAGGCCATCTT	609
Db	542	CGG--TTTCTGTGTCAGATGAGAGATCATTAACGAGTGGCAAGAGAG--CATCTT	599
QY	610	TTCTCATCGTTATTGTTCCTAGAGAGCTCTTCTGAGATC--TAGTTGGGCTTTCTTT	667

RESULT 7	
Bf674593	
LOCUS	BF674593
DEFINITION	Bf674593 902 bp mRNA linear EST 21-DEC-2001
ACCESSION	F02337811F NIH_MGC_03 Homo sapiens cDNA clone IMAGE:4274283 5',
VERSION	BF674593
KEYWORDS	BF674593.1 GI:11948488
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini;
	Homnidae; Homo.
REFERENCE	1 (bases 1 to 902)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-rcmail.nih.gov Tissue Procurement: CLONISTECH Laboratories, Inc. cDNA Library Preparation: CLONETECH Laboratories, Inc.

QY	12	GGGCAAAAATTAACAGGATATTGGTCCCGGAAACAAAGTTATCATTAAGATTAACAT	71
Db	2	GGGCAAAAATTAACAGGATATTGGTCCCGGAAACAAAGTTATCATTAAGATTAACAT	61
QY	72	TGATGCAGATGTTTCCGCCAAGCCACATATTTTCTTCAATTGTGTGAACCAAGCT	131
Db	62	TGATGCAGATGTTTCCGCCAAGCCACATATTTTCTTCAATTGTGTGAACCAAGCT	121
QY	132	CCAGAGGCTGGAACTACCTCTTGTCTTCTTGAGAAATTTTTCCTGATGTTATTAAGT	191
Db	122	CCAGAGGCTGGAACTACCTTGTCTTCTTGAGAAATTTTTCCTGATGTTATTAAGT	181
QY	192	ACATTGGCAAGAAAAGAGACAACGATTCGGGATCCGAGAGGGGAAACACCATGAA	251
Db	182	ACATTGGCAAGAAAAGAGACAACGATTCGGGATCCGAGAGGGGAAACACCATGAA	241
QY	252	GACTAACGACACATACATGAAATTTTAGCTGTTAAAGTGCCAGAAAATCACTGGACA	311
Db	242	GACTAACGACACATACATGAAATTTTAGCTGTTAAAGTGCCAGAAAATCACTGGACA	301
QY	312	AGAACACAGATGTATCGTCAGACATGANAATTAATTAACGAGTTGATCAAGAAATAT	371
Db	302	AGAACACAGATGTATCGTCAGACATGANAATTAATTAACGAGTTGATCAAGAAATAT	361
QY	372	CTTTCCTCCATTAAGACGAGTGTGATGACAAATGGATCCCAAGACAA--TTGTTCAAAAG	430
Db	362	CTTTCCTCCATTAAGACAGATGTGATGACAAATGGATCCCAAGACAAATTTGTTCAAAAG	421
QY	431	ATGCMAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTAACATGTAACCTCC	490
Db	422	ATGCMAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTAACATGTAACCTCC	481
QY	491	TCCGCTCCTCAAGAGTGGTCTATTTTGCCATCATACCGTGCTGTGCTTAAGAA	550
Db	482	TCCGCTCCTCAAGAGTGGTCTATTTTGCCATCATACCGTGCTGTGCTTAAGAA	541
QY	551	CGGCTTTCGTGTCAGATGAGAGAAAT--CATTAACGAGTGACCAAGAGCCATCTT	609
Db	542	CGG--TTTCGTGTCAGATGAGAGACATCATTAACGAGTGACCAAGAGG--CATCTT	599
QY	610	TTCTCATCGTTATTTGTCCTAGAAAGGCTCTTCTGAGATC--TAGTTGGGCTTTCTTT	667

Db 600 TTCCATCGGTATAGTCCCTAGAAAGCGTCTCTGAAGATCCAGGTGGGGCTCTCTT 659

Qy 668 CTGGGT-TTGGGCATTTCACTTCTCATGTC-TTACTATTTCTATCTATTATGATACG 725

Db 660 CTGGGTCTGGGGCCATTCAGTTCTCATGTGTGTTCATTCCTATTCATTTATGATTAAC 719

Qy 726 GTTTTCAACACAGTGGGCAC 745

Db 720 GGTTTAAACCAAGGCGCCAC 739

RESULT 8
AI685999/c
LOCUS
DEFINITION
638 bp mRNA linear EST 27-MAY-1999
t919104.x1 NCI CGAP Pr28 Homo sapiens CDNA clone IMAGE:2248878 3' similar to gb:M1321 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN); mRNA sequence.

ACCESSION
AI685999
EST.
AI685999.1 GI:4897293

VERSION
KEYWORDS
Homo sapiens (human)

SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS
TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT
Contact: Robert Straubeberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/INLW at: www-bio.lim.nih.gov/bdip/image/image.html
Seq primer: -40UP from G1Dco
High quality sequence stop: 453.

FEATURES
SOURCE
location/Qualifiers
1..638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2248878"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and 86 clones were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaudo."

ORIGIN
Query Match 59.8%; Score 613.8; DB 1; Length 638;
Best Local Similarity 98.6%; Pred. No. 1.2e-156;
Matches 629; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 391 GATGTCATCAACAGATCCCAAGACATTTTC-AAAGATCAATGATTAACAATCT 449

Db 638 GATGTCATCAACAGATCCCAAGACATTTTC-AAAGATCAATGATTAACAATCT 579

Qy 450 GCTGAGCTCAACAAACCTGTGATATTACATGACTCTCTGCTCCCAAGAGTGT 509

Db 578 GCTGAGCTCAACAAACCTGTGATATTACATGATCTCATCTCTCTCAAGAGTGT 519

Qy 510 GGTCTATTTTGCATATACACCTGCTGCTCTTAAGAAAGCGCTTTCGTCGAATGG 569

Db 518 GGTCTATTTTGCATATACACCTGCTGCTCTTAAGAAAGCGCTTTCGTCGAATGG 459

Qy 570 AGAAGATCATTAACAGAGCGTGGCACAAGAGCGCATTTTCTCATCGGTTATGTC 629

Db 458 AGAAGATCATTAACAGAGCGTGGCACAAGAGCGCATTTTCTCATCGGTTATGTC 399

Qy 630 CTAGAAAGCGTCTCTGAGATCTAGTGGCGTTCTTCTGCGTTTGGGCATTTCAATT 689

Db 398 CTAGAAAGCGTCTCTGAGATCTAGTGGCGTTCTTCTGCGTTTGGGCATTTCAATT 339

Qy 690 CTATGCTGTGATATTTCTATCATTTATTTGATTAACGCTTTTCAACAGTGGGCACAC 749

Db 338 CTATGCTGTGATATTTCTATCATTTATTTGATTAACGCTTTTCAACAGTGGGCACAC 279

Qy 750 AGAAGCTCACTCTGTATTAACATGAGAAATAGCCAGCGCATCTCCAGACCAATCTCT 809

Db 278 AGAAGCTCACTCTGTATTAACATGAGAAATAGCCAGCGCATCTCCAGACCAATCTCT 219

Qy 810 CCATGTTTTCACAGCTCTCTCCAGCCCAACCAATAGCGCTGCTATATGTAACATCC 869

Db 218 CCATGTTTTCACAGCTCTCTCCAGCCCAACCAATAGCGCTGCTATATGTAACATCC 159

Qy 870 TCGGCGTCTAGCGCTTGTCTCTCTTATGTTCTTTAATCAATATGCTCCGGAAGCC 929

Db 158 TCGGCGTCTAGCGCTTGTCTCTCTTATGTTCTTTAATCAATATGCTCCGGAAGCC 99

Qy 930 TTTCAATTTTACAGCGCTGAGAGATCTTCTGATGTAATTTATGTTGTTT 989

Db 98 TTTCAATTTTACAGCGCTGAGAGATCTTCTGATGTAATTTATGTTGTTT 39

Qy 990 CCGTAAATTAAGCAAAATTAATTTAAAAATGAAAAGTT 1027

Db 38 CCGTAAATTAAGCAAAATTAATTTAAAAATGAAAAGTT 1

RESULT 9
CN645427
LOCUS
DEFINITION
ILLUMIGEN MCO_23989 Katze_MSP Macaca mulatta cDNA clone
IBIDW:10243 5' similar to Bases 1 to 948 highly similar to human
TRG@ (Hs:385086), mRNA sequence.

ACCESSION
CN645427
EST.
CN645427.1 GI:47158870

VERSION
KEYWORDS
SOURCE
ORGANISM
Macaca mulatta (rhesus monkey)

REFERENCE
AUTHORS
Magnees, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B., Prohl, S.C., Fitzibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G., and Radonot, S.P.
Analysis of the Macaca mulatta transcriptome and the sequence divergence between Macaca and human
genome Biol. 6 (7), R60 (2005)

JOURNAL
PUBMED
15998449

COMMENT
Contact: C. Magnees
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagnees@illumigen.com
Sequenced on 2004.02.24. 794 Q20 bases.
PCR Primers
FORWARD: CCTCACTAAAGGAACAA
BACKWARD: CACTATAGGCGCAATGGGTA
Insert Length: 959 Std Error: 0.00

Plate: CL000136 row: G column: 03
Seq primer: CCTCCTCACTAAAGGAAACAAA
POLYA=yes.

FEATURES

source

Location/Qualifiers

1..959
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBIUM:10243"
/sex="male"
/cell_type="mononuclear lymphocyte"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_id="Katze_MMSP"
/note="Organ: spleen; Vector: Uni-ZAP XR; Site: 1; Site 2: Xho I; Created from Stragene ZAP-CDNA synthesis kit (catalog #200400) and ZAP-CDNA GigaPack III Gold Cloning Kit (catalog #200450)"

ORIGIN

Query Match 58.9%; Score 605.2; DB 7; Length 959;

Best Local Similarity 81.3%; Pred. No. 2.9e-154;
Matches 758; Conservative 0; Mismatches 78; Indels 96; Gaps 1;

189 GATACATTGGCAAGAAAGAGACACGATTTCTGGGATCCGAGAGGGGAAACACCAT 248
1 GATACATTGGCAAGAAAGAGACACGATTTCTGGGATCCGAGAGGGGAAACACCAT 60
249 GAAGCTAAGCAACATCATGAATTTAGCTGGTTACGGTGCAGAGAAAGTCACTGGA 308
61 GAAGCTAAGCAACATCATGAATTTAGCTGGTTACGGTGCAGAGAAAGTCACTGGA 120
309 CAAGAACAACAGATGATCTGTCAGACATGAGATTAATAAAGCGAGTTGATCAAGAAAT 368
121 TAAAGAGACAGATGATCTGTCAGACATGAGATTAATAAAGCGAGTTGATCAAGAAAT 180
369 TATCTTTCTCCCAATTAAGACGAGATG----- 395
181 TATCTTTCTCCCAATTAAGACAGATGTCACACAGTGGATCCCAAGACAGTTTTCAAA 240
396 ----- 395
241 AGACGCAATGATGTACACAGTGGATCCCAAGACAAATTAATTCAGAAATGCAGATGA 300
396 ---CATCAATGATGATCCCAAGACAAATGTTTCAAAAGATGCAATGATCACTACTGCT 452
301 TGACGCCACAGTGGATCCCAAGACAAATTAATTCAGAAATGCAGATGATGCAACTGCT 360
453 GCAGCTCAACAAACCTTGTGATTAATGATGATGCTCTCTCTCTCTCAAGAGTGGT 512
361 GCAGCTCAACAAACCTTGTGATTAATGATGATGCTCTCTCTCTCTCAAGAGTGGT 420
513 CTATTTTGGCATCAACCTGCTGCTGCTTGAAGAAAGCGCTTCTGCTCAATGAGA 572
421 CTATTTTGGCATCAACCTGCTGCTGCTTGAAGAAAGCGCTGCTGCTCAATGAGA 480
573 GAATCATTAACAGAGCGTGGCAACAAGAGCCATCTTTCTCATCGGTTATGTCCTTA 632
481 GAGATTTGAACAGAGCGTGGCAACAAGAGCCATCTTTCTCATCGGTTATGTCCTTA 540
633 GAAGGCTTTCTGAGGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCAATTCAGTTCTC 692
541 GAAGGCTTTCTGAGGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCAATTCAGTTCTC 600
693 ATGTGTGATCTATTCTATCATATTATTGTTAAACGGTTTTCAAACCGATGGGCAACAGAGA 752
601 ATGTGTGATCTATTCTATCATATTATTGTTAAAGTTTTCAAACCGATGGGCAACAGAGA 660
753 ACCTCACTCTGTAATTAACATGAGGATAGCAAGGCGCATCTCCAGCAACCATCTCTCCA 812
661 ACCTCACTCTGTAATTAACATGAGGATAGCAAGGCGCATCTCCAGCAACCATCTCTCCA 720

QY 813 TGTTTTCCAGCTCTCTCCAGCCCAACCAATAGCGCTGTATAGTGTAGACATCTCTGC 872
DB 721 TAGTCTCCAGAGCTCTCTCCAGCCCAACCAATAGCGCTGTATAGACATCTCTGC 780
QY 873 GGGTTAGCGCTGTCTCTCTCTAGTGTCTTTAATCAGATACTGCTGGAAGCCTTT 932
DB 781 AGCTTTAGCGCTGTCTCTCTCTAGTGTCTTTAATCAGATACTGCTGGAAGCCTTT 840
QY 933 CATTTTACAGCGCTGGAAGAGCTCTTGTGTAATGTAATGATGAGTGTGTTTCCG 992
DB 841 CATTTTACAGCTCTGGAAGAGCCTCTTGTGTAATGTAATGATGAGTGTGTTTCCG 900
QY 993 TAATAGCAAAATTAATTTAAAAATGAAA 1024
DB 901 AATAGCCAAATATATATTTAAAAATGAAA 932

RESULT 10

LOCUS

BC217853 703 bp mRNA linear EST 21-APR-2001
DEFINITION R5737574 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION

BC217853
VERSION BC217853.1 GI:13743874
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 703)
Harrington, J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramchandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McEligott, K., Booser, S.,
Mays, R., Smith, B., Veloso, N., Kliska, A., Hess, J., Cothren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL

PUBMED

COMMENT

Contact: Scott J. Cain
Atherys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherys.com
High quality sequence stop: 548.

FEATURES

source

Location/Qualifiers
1..703
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_id="HT1080"
/note="Atherys RAGE Library"

/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression', the
Nature Biotechnology', in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 57.5%; Score 590.2; DB 2; Length 703;

Best Local Similarity 97.3%; Pred. No. 3.5e-150;
Matches 621; Conservative 0; Mismatches 14; Indels 3; Gaps 2;

QY 391 GATGTATCAGATGATCCCAAGACAAATGTTTCAAAAGATGCAATGATGATCACTACTG 450
DB 53 GATGTATCAGATGATCCCAAGAC-ATTGGTCAAAAGATGCAATGATGATCACTACTG 111
QY 451 CTGAGCTCAACAAACACTCTGATATTATTCATGATGATGATGATGATGATGATGATG 510
DB 112 CTGAGCTCAACAAACACTCTGATATTATTCATGATGATGATGATGATGATGATG 171

QY 511 GTCTATTTTCATCATCATCCTGCTGCTCTTGTAGAGAAAGGCTTTCTGCTGCATATGA 570
DB 172 GTCTATTTTCATCATCATCCTGCTGCTCTTGTAGAGAAAGGCTTTCTGCTGCATATGA 231
QY 571 GAGAAATCATTAACAGACGCTGGGCAAGAGAGCCATCTTTTCTCATCGCTTATTTGCC 630
DB 232 GAGAAATCATTAACAGACGCTGGGCAAGAGAGCCATCTTTTCTCATCGCTTATTTGCC 291
QY 631 TAGAAGGCTTCTGAGGATCTAGTGGGCTTTCTTCTGGGTTTGGGCAATTTCAGTTTC 690
DB 292 TAGAAGGCTTCTGAGGATCTAGTGGGCTTTCTTCTGGGTTTGGGCAATTTCAGTTTC 351
QY 691 TCATGTGTACTATTTCTATCATTTATTTATTAACGCTTTTCAACACAGTGGGCAACAGA 750
DB 352 TCATGTGTACTATTTCTATCATTTATTTATTAACGCTTTTCAACACAGTGGGCAACAGA 411
QY 751 GAACCTCACTCTGTATTAATTAACATGAGAAATTAAGCCAGCTTCACACCAATCTCTC 810
DB 412 GAACCTCACTCTGTATTAATTAACATGAGAAATTAAGCCAGCTTCACACCAATCTCTC 471
QY 811 CATGTTTTCACAGCTCTCTCCAGCAACCCAAATAGGCTCTGCTATAGTATGACATCTC 870
DB 472 CATGTTTTCACAGCTCTCTCCAGCAACCCAAATAGGCTCTGCTATAGTATGACATCTC 531
QY 871 GCGGCTCTAGCTTGTCTCTCTCTTCTTATGATGATGATGATGATGATGATGATGATG 930
DB 532 GCGGCTCTAGCTTGTCTCTCTCTTCTTATGATGATGATGATGATGATGATGATGATG 591
QY 931 TTCAATTTTACAGCCCTGAAAGACGCTCTTCTGCTAGTGAATTAATGAG-1GTGTTT 988
DB 592 TTCAATTTTACAGCCCTGAAAGACGCTCTTCTGCTAGTGAATTAATGAGTGTGTTT 651
QY 989 TCCGTATTAAGCAAAATTAATTTAAAAAATGAAAAAT 1026
DB 652 CCGTATTAAGCAAAATTAATTTAAAAAATGAAAAAT 689

RESULT 11
A1768834/c 629 bp mRNA linear EST 21-DEC-1999
LOCUS A1768834
DEFINITION w03h04.x1 NCI CGAP K1d12 Homo sapiens CDNA clone IMAGE:2401783 3,
similar to gb:M3231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN);
mRNA sequence.

ACCESSION A1768834.1 GI:5235343
VERSION A1768834
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 629)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmer-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLM at:
www-bio.lnl.gov/bbrp/image/image.html
Insert length: 1658 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence, stop: 428.
Location/Qualifiers

FEATURES
SOURCE 1..629
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:2401783"
/feature_type="2 pooled tumore (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP K1d12"
/note="Organ: Kidney; Vector: pVT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP K1d5 was
prepared, and as circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match 57.4%; Score 589.6; DB 1; Length 629;
Best Local Similarity 97.9%; Pred. No. 5e-150;
Matches 606; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 406 GATCCCAAGACAAATTGTTCAAAAAGATGCAATGATGATGATGATGATGATGATGATG 465
DB 621 GATCCCAAGACAAATTG-TCAATGATGCAATGATGATGATGATGATGATGATGATG 563
QY 466 ACCTGCAATTAATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 525
DB 562 ACTGTCGATTAATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 503
QY 526 ATCACTGCTGCTGCTTGTAGAGAGAGCGCTTCTGCTGCAATGAGAGAGAGAGAGAG 585
DB 502 ATCACTGCTGCTGCTTGTAGAGAGAGCGCTTCTGCTGCAATGAGAGAGAGAGAGAG 443
QY 586 ACCGTGGACAG 645
DB 442 ACCGTGGACAG 383
QY 646 AGGATCTAGTGGGCTTTCTTCTGAGGTTTGGGCAATTCATGCTCATGCTGATGATG 705
DB 382 AGGATCTAGTGGGCTTTCTTCTGAGGTTTGGGCAATTCATGCTCATGCTGATGATG 323
QY 706 TCTATCATTAATTTATTAACCGTGTTCACACAGTGGGCAACAGAGAGAGAGAGAG 765
DB 322 TCTATCATTAATTTATTAACCGTGTTCACACAGTGGGCAACAGAGAGAGAGAGAG 263
QY 766 ATTAACAATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 825
DB 262 ATTAACAATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 203
QY 826 TCGTCCAGCAACCCAAATGAGGCTGCTATGATGATGATGATGATGATGATGATGATG 885
DB 202 TCGTCCAGCAACCCAAATGAGGCTGCTATGATGATGATGATGATGATGATGATGATG 143
QY 886 GTCCCTCTCTTATGTTCTTATTAATCAGATGATGATGATGATGATGATGATGATGATG 945
DB 142 GTCCCTCTCTTATGTTCTTATTAATCAGATGATGATGATGATGATGATGATGATGATG 83
QY 946 CTGAGCAGTCTTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1005
DB 82 CTGAGCAGTCTTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 23
QY 1006 AATTTTAAAAAATGAAAA 1024
DB 22 AATTTTAAAAAATGAAAA 4

RESULT 12
CA449324/c 700 bp mRNA linear EST 08-NOV-2002
LOCUS CA449324
DEFINITION UT-H-E11-ayc-k-17-0-UT.s1 NCI CGAP E11 Homo sapiens CDNA clone
ACCESSION CA449324
VERSION CA449324.1 GI:24813744

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 700)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-29, >AT rich#low_complexity
Seq primer: M13 FORWARD
POLY-A=yes

FEATURES
Source
1..700
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-B11-ayc-k-17-0-UI"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP_B11 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ACACTGGCAC
TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-B11
TAG_SEQ=ACACTGGCAC"

ORIGIN
Query Match 56.9%; Score 584; DB 6; Length 700;
Beet local Similarity 95.9%; Pred. No. 1.7e-148; Indels 22; Gaps 2;
Matches 629; Conservative 0; Mismatches 5;
391 GATGTCATCAATGATCCAAAGACATTTTCAAAAGATGCAAT----- 438
661 GATGTCATCAATGATCCAAAGACATTTGCTAAAGATGCAATGATGCAA 602
439 -----GATACCTACTGCTGCACTCACAACACCTCTGATATTACATGTACTC 489
601 TTTCTCTTAATACCTACTGCTGCACTCACAACACCTCTGATATTACATGTACTC 542
490 CTCCTGCTCTTC-AGAGTGTGTCTATTGTCCTATCCTGCTCTGCTTGAAG 548
541 CTCCTGCTCTTCMAAGAGTGTGTCTATTGTCCTATCCTGCTCTGCTTGAAG 482
549 AACGGCTTTGCTGCTCAATGAGAGAAATCATTAACAGACGGTGCACAAGAGGCATCT 608
481 AACGGCTTTGCTGCTCAATGAGAGAAATCATTAACAGACGGTGCACAAGAGGCATCT 422
609 TTTCTCATCGGTATTGTCTCCTAGAACGCTTTCTGAGATCTAGTTGGGCTTTCTTTC 668

Db 421 TTTCTCATCGGTATTGTCTCCTAGAACGCTCTTCTGAGGATCTAGTGGGCTTTCTTC 362
Qy 669 TGGGTTTGGGCAATTCAGTTCTCATGTGTACTATTATCATTAATTAACGGTT 728
Db 361 TGGGTTTGGGCAATTCAGTTCTCATGTGTACTATTATCATTAATTAATGTT 302
Qy 729 TTCAACACGATGGGCAACAGAGAACTCACTCTGTATAACATGAGAAATGCGCAG 788
Db 301 TTCAACACGATGGGCAACAGAGAACTCACTCTGTATAACATGAGAAATGCGCAG 242
Qy 789 CGATCTCCAGCAACCAATCTCATATGTTTCCACAGCTCTCCAGCCCAACCAATGAGC 848
Db 241 CGATCTCCAGCAACCAATCTCATATGTTTCCACAGCTCTCCAGCCCAACCAATGAGC 182
Qy 849 CTTGCTATGATGATACATCTGCGGCTTCTAGCTTCTCTCTTCTTCTTCTTAA 908
Db 181 CTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
Qy 909 TCAATTAATGCTGCTGGAAGCTTTTCAATTTTACACGCCCTGAAGCAGTCTTCTTCTAGT 968
Db 121 TCAATTAATGCTGCTGGAAGCTTTTCAATTTTACACGCCCTGAAGCAGTCTTCTTCTAGT 62
Qy 969 TCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
Db 61 TCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6

RESULT 13
LOCUS BE326754/c 582 bp mRNA linear EST 14-JUN-2000
DEFINITION h64d05.x1 NCI CGAP_Kid11 Homo sapiens cDNA clone IMAGE:313257 3'
similar to gb:M13231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN);,
mRNA sequence.
ACCESSION BE326754 GI:9200530
VERSION BE326754
KEYWORDS EST.
ORGANISM Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 582)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
Clone Distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gldco
High quality sequence stop: 447.
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:313257"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
plasmid DNA from the normalized library NCI CGAP_Kid11 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subcloning
hybridization reaction. The driver was PCR-amplified cDNA

ORIGIN

from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo.

Query Match 56.5%; Score 580.4; DB 2; Length 582;
Best Local Similarity 99.8%; Pred. No. 1.6e-147;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 439 GATACACTACTGCTGCGCTCACAACAACCTTCGATATTAACATGATCTCTCTGCTC 438
Db 582 GATACACTACTGCTGCGCTCACAACAACCTTCGATATTAACATGATCTCTCTGCTC 523
Qy 499 CTCAAGAGTGTGCTATTTTGGCATCATCACTGCTGCTGTAGAAAGAGGCTTTC 558
Db 522 CTCAAGAGTGTGCTATTTTGGCATCATCACTGCTGCTGTAGAAAGAGGCTTTC 463
Qy 559 TGCTGCATGAGAGAAATCATACAGACGGTGGCACAAGAGGCACTCTTTCCTCATC 618
Db 462 TGCTGCATGAGAGAAATCATACAGACGGTGGCACAAGAGGCACTCTTTCCTCATC 403
Qy 619 GGTATTTGCTCCAGAAAGCGTCTTCTGAGATCTAGTTGGGCTTCTTCTGCGTTGGG 678
Db 402 GGTATTTGCTCCAGAAAGCGTCTTCTGAGATCTAGTTGGGCTTCTTCTGCGTTGGG 343
Qy 679 CCATTTGAGTCTGATGCTGCTATTTATTCATTAATTTGATTAACGGTTTCAACAG 738
Db 342 CCATTTGAGTCTGATGCTGCTATTTATTCATTAATTTGATTAACGGTTTCAACAG 283
Qy 739 TGGGCAACAGAGAACTCTGATTAACAATGAGAAATAGCAGCGATCTCAG 798
Db 282 TGGGCAACAGAGAACTCTGATTAACAATGAGAAATAGCAGCGATCTCAG 223
Qy 799 CACCAATCTCTCATGTTTTCACAGCTCTCCAGCCACCAATAAGCGCTGCTAG 858
Db 222 CACCAATCTCTCATGTTTTCACAGCTCTCCAGCCACCAATAAGCGCTGCTAG 163
Qy 859 TGTGACATCTGCGGCTTCTAGGCTTGTCCCTCTTATGTTCTTATACATTAAT 918
Db 162 TGTGACATCTGCGGCTTCTAGGCTTGTCCCTCTTATGTTCTTATACATTAAT 103
Qy 919 GCCTGAAGCCTTTCATTTTACAGCCCTGAAAGAGCTTCTTGTAGTTGAATTAAT 978
Db 102 GCCTGAAGCCTTTCATTTTACAGCCCTGAAAGAGCTTCTTGTAGTTGAATTAAT 43
Qy 979 GGTGTGTTTTCCTTAATAAGCAAAATTAATTTAAAAAATG 1020
Db 42 GGTGTGTTTTCCTTAATAAGCAAAATTAATTTAAAAAATG 1

RESULT 14 613 bp mRNA linear EST 16-OCT-2001
BI905966 60306285F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212261 5',
LOCUS mRNA sequence.
ACCESSION BI905966
VERSION BI905966.1 GI:16168605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 613)
NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
AUTHORS Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LLM11532 row: 0 column: 14
High quality sequence stop: 578.
Location/Qualifiers
1. 613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5212261"
/cissue_type="Leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source Leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Query Match 56.4%; Score 578.8; DB 3; Length 613;
Best Local Similarity 99.7%; Pred. No. 4.5e-147;
Matches 580; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 443 CACTACTGCTGACGCTCACAAACCTCTGCATATTATACATGATCTCTCTCTCTCA 502
Db 1 CACTACTGCTGACGCTCACAAACCTCTGCATATTATACATGATCTCTCTCTCTCA 60
Qy 503 AGAGTGTGCTATTTTGGCATATCATCCTGCTGTCTGCTTAAAGAAAGCGCTTCTGCT 562
Db 61 AGAGTGTGCTATTTTGGCATATCATCCTGCTGTCTGCTTAAAGAAAGCGCTTCTGCT 120
Qy 563 GCAATGAGAGAAATCATTAACAGCGTGGCAACAGAGGCACTTTTCTCATCGGTT 622
Db 121 GCAATGAGAGAAATCATTAACAGCGTGGCAACAGAGGCACTTTTCTCATCGGTT 180
Qy 623 ATTGTCCTTAAGAGCTCTTCTAGAGATCTAGTTGGGCTTCTTCTGCTTGGGCAAT 682
Db 181 ATTGTCCTTAAGAGCTCTTCTAGAGATCTAGTTGGGCTTCTTCTGCTTGGGCAAT 240
Qy 683 TTCAAGTTCATGTTGCTATTTATCATTAATTTGATTAACGATTTTCAACCAAGTGG 742
Db 241 TTCAAGTTCATGTTGCTATTTATCATTAATTTGATTAACGATTTTCAACCAAGTGG 300
Qy 743 CACACAGAAACCTCACTCTGTAATAACAATGAGAAATACCAAGGATCTCCAGACC 802
Db 301 CACACAGAAACCTCACTCTGTAATAACAATGAGAAATACCAAGGATCTCCAGACC 360
Qy 803 AATCTGCTCAATTTTCCACAGCTCTCCAGCAACCAATAGAGGCTCTATAGTGA 862
Db 361 AATCTGCTCAATTTTCCACAGCTCTCCAGCAACCAATAGAGGCTCTATAGTGA 420
Qy 863 GACATCTGCGGCTTCTAGGCTTGTCCCTCTCTTATGTTCTTATACATTAATCTGCT 922
Db 421 GACATCTGCGGCTTCTAGGCTTGTCCCTCTCTTATGTTCTTATACATTAATCTGCT 480
Qy 923 GGAAGCTTTCATTTTACAGCCCTGAAAGAGCTTCTTGTGTAATTTATGTTGGTG 982
Db 481 GGAAGCTTTCATTTTACAGCCCTGAAAGAGCTTCTTGTGTAATTTATGTTGGTG 540
Qy 983 TGTATTTCCGTAATAAGCAAAATTAATTTAAAAAATGAATA 1024
Db 541 TGTATTTCCGTAATAAGCAAAATTAATTTAAAAAATGAATA 582

RESULT 15
BF681238

LOCUS BF681238 719 bp mRNA linear EST 21-DEC-2000
 DEFINITION 602155502F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4296504 5',
 mRNA sequence.
 ACCESSION BF681238
 VERSION BF681238.1 GI:11955133
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 719)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LUCM1149 row: c column: 01
 High quality sequence stop: 673.
 Location/Qualifiers
 1..719
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4296504"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 83"
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
 Site_1: SfiI (ggcgcctcgcc); Site_2: SfiI
 (ggccatcaggcc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGCGCGGCACATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA)."

ORIGIN

Query Match 55.5%; Score 569.8; DB 2; Length 719;
 Best Local Similarity 94.6%; Pred. No. 1.4e-144;
 Matches 678; Conservative 0; Mismatches 27; Indels 12; Gaps 8;

```

OY 11 TGGGCAAAAAATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAAC 70
    |||
DB 5 TGGGCAAAAAATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAAC 64
    |||
OY 71 TTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCCTCAATGCTGAAACAAGC 130
    |||
DB 65 TTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCCTCAATGCTGAAACAAGC 124
    |||
OY 131 TCCAGAGGCTGGAAATACCTTTGTCTTTGAGAAATTTTCCCTGATGTTATTAGA 190
    |||
DB 125 TCCAGAGGCTGGAAATACCTTTGTCTTTGAGAAATTTTCCCTGATGTTATTAGA 183
    |||
OY 191 TACATTGGCAAGAAAGAGAGCAACAGATTCTGGGATCCCAAGAGGGGAAACACCATGA 250
    |||
DB 184 TACATTGGCAAGAAAGAGAGCAACAGATTCTGGGATCCCAAGAGGGGAAACACCATGA 243
    |||
OY 251 AGACTTAAGACACATACATGAATTTAGTGTGTTAACGGTGCAGAAAAAGTCACTGACA 310
    |||
DB 244 AGACTTAAGACACATACATGAATTTAGTGTGTTAACGGTGCAGAAAAAGTCACTGACA 303
    |||
OY 311 AAGAACACAGATGTTCTGTCAGATGAGATTAATMAAACGAGTTGATCAAGAAATTA 370
    |||

```

```

DB 304 AAGAACACAGATGTTCTGTCAGATGAGATTAATMAAACGAGTTGATCAAGAAATTA 363
OY 371 TCTTCTCCCAATPAAGACGAGATGTCACAATGATCCCAAGACAAATTTTCAAAAG 430
    |||
DB 364 TCTTCTCCCAATPAAGACGAGATGTCACAATGATCCCAAGACAAATTTTCAAAAG 423
    |||
OY 431 ATGCAATGATACACTACTGCTGAGCTCAAAACACTCTGCAATTTACATGTAACCTCC 490
    |||
DB 424 ATGCAATGATACACTACTGCTGAGCTCAAAACACTCTGCAATTTACATGTAACCTCC 483
    |||
OY 491 TCTGCTCTCAAGAGTGTGTTA--TTTGCCATCATCACTGCTGCTT--AGAAAG 548
    |||
DB 484 TCTGCTCTCAAGAGTGTGTTA--TTTGCCATCATCACTGCTGCTT--AGAAAG 543
    |||
OY 549 AACGGCTTTCTG--CTGCAATGAGAGAAATCATPAACAGACGTGCAAC--AAGAGGCCA 605
    |||
DB 544 AACGGCTTTCTGCGTGAATTGAGAGAAATCATPAACAGACGTGCAAC--AAGAGGCCA 603
    |||
OY 606 TCTTTTCTCATCGGTTA--TTGTCCCTAGAAAGGCTCTTC--TGAGATCTAGTTGGGCTT 662
    |||
DB 604 TCTTTTCTCATCGGTTA--TTGTCCCTAGAAAGGCTCTTC--TGAGATCTAGTTGGGCTT 663
    |||
OY 663 TCTTTCTGAGTTGGGCAATTCAGTTCTCATGTTGTTACTATCTATCATTTATGT 719
    |||
DB 664 ACTTCTGGGTTGGCAATTCAGTTCTCATGTTGTTACTATCTATCATTTATGT 717
    |||

```

Search completed: December 10, 2005, 23:23:07
 Job time : 4086 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 17:49:47 ; Search time 664 Seconds
(without alignments)
10308.192 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027
Sequence: 1 gggcagagcttggcacaataa.....attcaaaaaatgcaaacgtc 1027

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_21:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1027	100.0	1027	4	AAFS6401 Human TCR
2	1027	100.0	1027	4	ADB75580 Prostate
3	1027	100.0	1027	13	ADP54862 Human PRO
4	1027	100.0	1027	14	ADM4400 Human TAR
5	1013.8	98.7	1395	12	ADL06443 Human tum
6	984	95.8	1530	14	ADY16296 DNA encod
7	968.2	94.3	1799	6	ABZ35381 Human gen
8	945.6	92.1	1418	6	ABD38828 Human PSN
9	915	89.1	2658	10	ADBE08798 Novel DNA
10	906	88.2	1155	3	AAFI5623 Human pro
11	905.6	88.2	1560	10	ADBE06987 Novel cod
12	902.2	87.8	1421	6	ABL65007 lung canc
13	902.2	87.8	1421	13	ADRC5880 Human pro
14	902.2	87.8	1421	13	ADRC6783 Human pro
15	838	81.6	1586	13	ABK64529 Human ben
16	838	81.6	1586	13	ADP54939 Human PRO
17	838	81.6	1586	14	ACI38957 TRG codin
18	838	81.6	1586	14	ADY16545 DNA encod
19	835.2	81.3	2658	10	ADBE08798 Novel DNA

20	821.6	80.0	1046	12	ADP10449 Reference
21	795.8	77.5	1162	5	ABV27727 Human pro
22	795.8	77.5	1162	5	ABV24153 Human pro
23	795.8	77.5	1162	5	ABV24540 Human pro
24	795.8	77.5	1162	5	ABV25162 Human pro
25	795.8	77.5	1162	5	ABV21526 Human pro
26	795.8	77.5	1162	5	ABV27345 Human pro
27	795.8	77.5	1162	5	ABV24544 Human pro
28	795.8	77.5	1162	5	ABV24851 Human pro
29	660.2	64.3	1665	14	ACI55249 Human col
30	619	60.3	924	5	ABV25092 Human pro
31	601	58.5	786	13	ADP55163 Human PRO
32	589	57.4	124933	14	ABE61187 Human TRG
33	582.6	56.7	765	14	ADY78562 Human CDV
34	522.8	50.9	1080	14	AAO37617 TCR gamma
35	522.8	50.9	1080	14	ADY20108 DNA encod
36	516.4	50.3	1080	1	AAAN1698 Sequence
37	498.8	48.6	1316	5	ABV26485 Human pro
38	498.8	48.6	1316	5	ABV25093 Human pro
39	492	47.9	531	13	ADU12452 Solid tum
40	473.2	46.1	486	5	ABV25013 Human pro
41	449.2	43.7	825	2	AAQ66888 Human lym
42	449.2	43.7	825	9	ACF35988 Human T c
43	446.8	43.5	784	5	ABV21896 Human pro
44	444	43.2	511	2	AAV58549 Prostate
45	444	43.2	511	2	AAV61291 CDNA sequ

ALIGNMENTS

RESULT 1
AAFS6401
ID AAF56401 standard; cDNA; 1027 BP.
XX AAF56401;
AC
XX
DT 12-APR-2001 (first entry)
XX
DE Human TCRgamma alternate reading frame protein coding sequence.
XX
KM Human; TARP; prostate cancer; breast cancer; immunotherapy;
KM T cell receptor gamma alternate reading frame protein; TCRgamma; ss.
OS Homo sapiens.
XX
PN WO200104309-A1.
XX
PD 18-JAN-2001.
XX
PF 12-JUL-2000; 2000NC-US019039.
XX
PR 13-JUL-1999; 99US-0143560P.
PR 01-OCT-1999; 99US-0157471P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pastan I, Essand M, Lee B, Vasmatzis G, Wolfgang C;
XX
DR WPI: 2001-081050/09.
XX P-PsDB; AAB66399, AAB66400.
XX
PT Isolated T-cell receptor gamma alternate reading frame protein useful for
PT diagnosing and raising an immune response to prostate cancer and breast
XX cancer.
XX
PS Example 1; Fig 1; 85pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human T cell receptor alternate reading frame protein (TARP). This
CC protein is expressed in prostate and breast cancer cells at higher levels
CC than normal and so can be used in the immunotherapy of these cancers, as
CC well as their detection and prevention


```

XX Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;
SQ
Query Match      100.0%; Score 1027; DB 4; Length 1027;
Best Local Similarity 100.0%; Pred. No. 4.3e-290;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTAACA 60
DB 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTAACA 60
QY 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120
DB 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120
QY 121 GAAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGAT 180
DB 121 GAAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGAT 180
QY 181 GTTATTTAGATACATTTGGCAAGAAAGAGCAACAGATTTGCGATCCGAGAGGGG 240
DB 181 GTTATTTAGATACATTTGGCAAGAAAGAGCAACAGATTTGCGATCCGAGAGGGG 240
QY 241 AACACCATGAAGACTAAGACACATACATGAAATTTAGCTGTTAACGTCGCAAGAAAG 300
DB 241 AACACCATGAAGACTAAGACACATACATGAAATTTAGCTGTTAACGTCGCAAGAAAG 300
QY 301 TCACCTGGCAAGAAACACAGATGTATCGTCAGACATGAGAAATTAATTAAGAGGAGTTGAT 360
DB 301 TCACCTGGCAAGAAACACAGATGTATCGTCAGACATGAGAAATTAATTAAGAGGAGTTGAT 360
QY 361 CAAGAAATTTATCTTCTCTCAATAAAGAGATGTCATACAGATGATCCCAAGACAT 420
DB 361 CAAGAAATTTATCTTCTCTCAATAAAGAGATGTCATACAGATGATCCCAAGACAT 420
QY 421 TGTTCAAAAAGATGCAATGATACACTACTGCTGACAGACACCTCTGCATATTAC 480
DB 421 TGTTCAAAAAGATGCAATGATACACTACTGCTGACAGACACCTCTGCATATTAC 480
QY 481 ATGTACCTCTCTGCTCTCTCAAGAGTGGTCTATTTTGGCATCATGACCGCTGCTG 540
DB 481 ATGTACCTCTCTGCTCTCTCAAGAGTGGTCTATTTTGGCATCATGACCGCTGCTG 540
QY 541 CTTAGAAAGACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAACAAGA 600
DB 541 CTTAGAAAGACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAACAAGA 600
QY 601 GGGCATCTTTTCTCTATCGGTTATTTGCTTGAAGCGCTTTCTGAGATCTAGTTGGGC 660
DB 601 GGGCATCTTTTCTCTATCGGTTATTTGCTTGAAGCGCTTTCTGAGATCTAGTTGGGC 660
QY 661 TTTCTTCTGGGTTGGGCAATTTAGTTCTCATGTTGTAATCTATCTATCATTTATGTA 720
DB 661 TTTCTTCTGGGTTGGGCAATTTAGTTCTCATGTTGTAATCTATCTATCATTTATGTA 720
QY 721 TAAACGTTTTCACCAAGTGGGACACAGAGAACTCACTCTGTATTAACAATGAGAT 780
DB 721 TAAACGTTTTCACCAAGTGGGACACAGAGAACTCACTCTGTATTAACAATGAGAT 780
QY 781 AGCCAGCGGATCTCCAGACCAATCTCTCCATGTTTTCACAGCTCTCCGACCAACC 840
DB 781 AGCCAGCGGATCTCCAGACCAATCTCTCCATGTTTTCACAGCTCTCCGACCAACC 840
QY 841 AAATAGCGCTGTATGATAGATCTCGGGCTTGAAGCTTGCCTCTCTTAGTG 900
DB 841 AAATAGCGCTGTATGATAGATCTCGGGCTTGAAGCTTGCCTCTCTCTTAGTG 900
QY 901 TTCTTTAATACAGATACTGCTGAGAGCTTTCACTTTTACAGCCCTGAGAGAGCTTCT 960
DB 901 TTCTTTAATACAGATACTGCTGAGAGCTTTCACTTTTACAGCCCTGAGAGAGCTTCT 960
QY 961 TTGCTAGTTGAATTAATGAGTGTGTTTTCGTAATTAAGCAAAATTAATTAATAATG 1020
DB 961 TTGCTAGTTGAATTAATGAGTGTGTTTTCGTAATTAAGCAAAATTAATTAATAATG 1020

```

```

DB 961 TTGCTAGTTGAATTAATGAGTGTGTTTTCGTAATTAAGCAAAATTAATTAATAATG 1020
QY 1021 AAAAGTT 1027
DB 1021 AAAAGTT 1027

RESULT 2
ADB75580
ID ADB75580 standard; cDNA; 1027 BP.
XX
AC ADB75580;
XX
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker cDNA.
XX
KM Prostate; cancer; cytostatic; gene therapy; marker; ss.
XX
OS Homo sapiens.
XX
PN M02003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
PR 22-AUG-2001; 2001US-0314356P.
PR 25-SEP-2001; 2001US-0325020P.
PR 12-DEC-2001; 2001US-0341746P.
PR 05-MAR-2002; 2002US-0362158P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Monahan JF, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoerish S, Kamatkar S, Monsey AM, Glatt K, Zhao X, Anderson D;
XX
DR WPI; 2003-248033/24.
XX
PT New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX
PS Disclosure; SEQ ID NO 404; 99pp; English.
XX
CC The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;
Query Match      100.0%; Score 1027; DB 10; Length 1027;
Best Local Similarity 100.0%; Pred. No. 4.3e-290; Indels 0; Gaps 0;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTAACA 60
DB 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTAACA 60
QY 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120
DB 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120

```


Db 1 GGGCAGAGTTGGGCAAAAAATTAAGTAATTGGTCCGAGACAAAGCTTATCATTTACA 60
 Qy 61 GATTAACAACTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTTGCT 120
 Db 61 GATTAACAACTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTTGCT 120
 Qy 121 GAAACAAAGCTCCGAAAGGCTGGAACATACCTTTGTCTTCTTGAAATTTTCCCTGAT 180
 Db 121 GAAACAAAGCTCCGAAAGGCTGGAACATACCTTTGTCTTCTTGAAATTTTCCCTGAT 180
 Qy 181 GTTATTAAATATCACTTGGCAAGAAAAGAGCAACAGATTCGGGATCCGAGAGGG 240
 Db 181 GTTATTAAATATCACTTGGCAAGAAAAGAGCAACAGATTCGGGATCCGAGAGGG 240
 Qy 241 AACACATGAAGACTAAAGACATACATGAAATTTAGCTGTTACGGTGCCAGAAAG 300
 Db 241 AACACATGAAGACTAAAGACATACATGAAATTTAGCTGTTACGGTGCCAGAAAG 300
 Qy 301 TCACATGACAAAGAACACAGATGATGCTCAGACATGAAATTAATAAAACGAGTTGAT 360
 Db 301 TCACATGACAAAGAACACAGATGATGCTCAGACATGAAATTAATAAAACGAGTTGAT 360
 Qy 361 CAAGAAATTAATCTTCCCAATTAAGACGGATGTCATCAATGATCCCAAGAGCAAT 420
 Db 361 CAAGAAATTAATCTTCCCAATTAAGACGGATGTCATCAATGATCCCAAGAGCAAT 420
 Qy 421 TGTTCAAAGATGCAAAATGATACACTACTGCTGAGCTCAAAACACCTGCAATTTAC 480
 Db 421 TGTTCAAAGATGCAAAATGATACACTACTGCTGAGCTCAAAACACCTGCAATTTAC 480
 Qy 481 ATGTACTCTCTCTGCTCTCAGAGTGTGTCATTTTGGCATATCACTGCTGCTG 540
 Db 481 ATGTACTCTCTCTGCTCTCAGAGTGTGTCATTTTGGCATATCACTGCTGCTG 540
 Qy 541 CTTAGAGAAACGGCTTCTGCTGCAATGAGAAATCATTAACAGAGTGGGCAAGAGA 600
 Db 541 CTTAGAGAAACGGCTTCTGCTGCAATGAGAAATCATTAACAGAGTGGGCAAGAGA 600
 Qy 601 GGCCATCTTTTCTCATCGGTTATTTGTCCTTGAAGAGGCTTCTGAGAGATCTAGTGGC 660
 Db 601 GGCCATCTTTTCTCATCGGTTATTTGTCCTTGAAGAGGCTTCTGAGAGATCTAGTGGC 660
 Qy 661 TTTCTTCTGGGTTTGGGCCATTTCAAGTCTCATGTTGTAATTTCTATATTATTGTA 720
 Db 661 TTTCTTCTGGGTTTGGGCCATTTCAAGTCTCATGTTGTAATTTCTATATTATTGTA 720
 Qy 721 TTAAGGTTTCAAAACAGTGGGCAACAGAGAACCTCATGTAATTAACAAATGAGGAT 780
 Db 721 TTAAGGTTTCAAAACAGTGGGCAACAGAGAACCTCATGTAATTAACAAATGAGGAT 780
 Qy 781 AGCCACGGCGATCTCCAGACCAATCTCTCCATGTTTTCACACAGCTCTCCAGCAACC 840
 Db 781 AGCCACGGCGATCTCCAGACCAATCTCTCCATGTTTTCACACAGCTCTCCAGCAACC 840
 Qy 841 AAATAGGGCTGCTAATGATGTAACATCTCGGGCTTACGCTTGTCTCTTATG 900
 Db 841 AAATAGGGCTGCTAATGATGTAACATCTCGGGCTTACGCTTGTCTCTTATG 900
 Qy 901 TTTCTTATCAGATTAAGCTGGAAGCTTCAATTTTACAGGCGCTGGAAGCACTTTCT 960
 Db 901 TTTCTTATCAGATTAAGCTGGAAGCTTCAATTTTACAGGCGCTGGAAGCACTTTCT 960
 Qy 961 TTGCTAGTTGAATTAATGATGTTTCCGTAATAAGCAAAATTAATTAATAAATG 1020
 Db 961 TTGCTAGTTGAATTAATGATGTTTCCGTAATAAGCAAAATTAATTAATAAATG 1020
 Qy 1021 AAAAGTT 1027
 Db 1021 AAAAGTT 1027

RESULT 4
 ADM44400

ID ADM44400 standard; DNA; 1027 BP.
 AC ADM44400;
 XX
 XX
 DT 24-MAR-2005 (first entry)
 XX
 XX
 DE Human TARP polynucleotide.
 XX
 XX T-cell receptor gamma alternate reading frame protein; TARP;
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm; gene;
 KM ds.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 74..250
 FT /*tag= a
 FT /product= "Human TARP"
 PN WC200500089-A1.
 PD 06-JAN-2005.
 PF 02-JUN-2004; 2004MO-US017574.
 XX
 XX 05-JUN-2003; 2003US-0476467P.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Berzofsky JA, Oh S, Pastan I;
 XX WPI; 2005-091494/10.
 DR P-PSDB; ADM44399.
 DR GENBANK; AF151103.
 XX
 PT New T-cell receptor gamma alternate reading frame protein (TARP)
 PT polypeptide, useful for treating a subject having breast or prostate
 PT cancer and for generating an immune response to TARP-expressing breast
 PT and prostate cancer cells.
 XX
 PS Disclosure; SEQ ID NO 2; 83pp; English.
 XX
 CC The invention relates to an immunogenic T-cell receptor gamma alternate
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding
 CC it. The invention also relates to a vector comprising the TARP
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical
 CC composition comprising a therapeutic amount of the polypeptide or the
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response
 CC in a subject, inhibiting the growth of a breast cancer or a prostate
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
 CC detecting T cells expressing CD8 that specifically recognize the TARP
 CC polypeptide in a subject. Eliciting an immune response in a subject
 CC comprises administering the polypeptide or polynucleotide, thus producing
 CC an immune response in the subject. The immune response comprises a T cell
 CC response or inducing cytotoxic T cells that induce lysis of cells
 CC expressing the TARP polypeptide. The subject has breast cancer or
 CC prostate cancer. The immune response decreases the growth of the prostate
 CC cancer or breast cancer. The method further comprises administering an
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
 CC or CTL precursor cells with the polypeptide and an antigen presenting
 CC cell to produce activated CTLs matured from the CTL precursors that
 CC recognize the breast cancer or the prostate cancer cells, and contacting
 CC the breast cancer or the prostate cancer cell with the activated CTLs or
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and
 CC polynucleotide are useful for treating a subject having breast or
 CC prostate cancer. The polypeptide is useful for generating an immune
 CC response to breast cancer and prostate cancer cells that express TARP
 CC polypeptides. This sequence represents the human TARP polynucleotide of
 CC the invention.
 XX

Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;

Query Match 100.0%; Score 1027; DB 14; Length 1027;
Best Local Similarity 100.0%; Pred. No. 4.3e-290;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGCAAGTGGGCAAAAATCAAGTATTTGGTCCCGCAACAAAGCTTATCATTA 60
DB 1 GGGCAAGTGGGCAAAAATCAAGTATTTGGTCCCGCAACAAAGCTTATCATTA 60
QY 61 GATTAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCT 120
DB 61 GATTAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCT 120
QY 121 GAAACAAGCTCCGAAAGGCTGGAACATCTTGTCTCTTGAGAAATTTTCCCTGAT 180
DB 121 GAAACAAGCTCCGAAAGGCTGGAACATCTTGTCTCTTGAGAAATTTTCCCTGAT 180
QY 181 GTTATTAAGATACATTTGGCAAGAAAGAGCAACAGATTCTGGATCCCAAGAGGG 240
DB 181 GTTATTAAGATACATTTGGCAAGAAAGAGCAACAGATTCTGGATCCCAAGAGGG 240
QY 241 AACACATGAAGACTAAAGACATACATGAAATTTAGCTGTACGGTGCAGAAAG 300
DB 241 AACACATGAAGACTAAAGACATACATGAAATTTAGCTGTACGGTGCAGAAAG 300
QY 301 TCATCGACAAAGAAACACAGTGTATGCTCAGACATGAAATTAATAAAACGGATGAT 360
DB 301 TCATCGACAAAGAAACACAGTGTATGCTCAGACATGAAATTAATAAAACGGATGAT 360
QY 361 CAAGAAATTTCTTCCCTCCAAATAAGACGAGTGCATCAATGATGCCAAGACAT 420
DB 361 CAAGAAATTTCTTCCCTCCAAATAAGACGAGTGCATCAATGATGCCAAGACAT 420
QY 421 TGTTCAAAAGATGCAAAATGATACATCTGCTGACGCTCAACAAACCTCTGATATTAC 480
DB 421 TGTTCAAAAGATGCAAAATGATACATCTGCTGACGCTCAACAAACCTCTGATATTAC 480
QY 481 ATGTACCTCTCTGCTGCTCTCAAGATGTGTGATTTTGGCATCATCACTGCTGTCTG 540
DB 481 ATGTACCTCTCTGCTGCTCTCAAGATGTGTGATTTTGGCATCATCACTGCTGTCTG 540
QY 541 CTTAGAAAGACGGCTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGAGA 600
DB 541 CTTAGAAAGACGGCTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGAGA 600
QY 601 GGCCATCTTTTCTCATCGGTTATGTCCTTAGAAGGCTCTTCTGAGGATCTAGTGGGC 660
DB 601 GGCCATCTTTTCTCATCGGTTATGTCCTTAGAAGGCTCTTCTGAGGATCTAGTGGGC 660
QY 661 TTTCTTTCTGGGTTGGGCCATTTCAAGTTCATGATGTGATCTATTCATATTGTA 720
DB 661 TTTCTTTCTGGGTTGGGCCATTTCAAGTTCATGATGTGATCTATTCATATTGTA 720
QY 721 TTAAGGTTTAAACCAAGTGGGCAACAGACACCTCACTGATTAACAATAGGAGAT 780
DB 721 TTAAGGTTTAAACCAAGTGGGCAACAGACACCTCACTGATTAACAATAGGAGAT 780
QY 781 AGCCACGGCGATCTCTCAAGCAAAATCTCTCATGTTTTCACAGCTCTCCAGCAACC 840
DB 781 AGCCACGGCGATCTCTCAAGCAAAATCTCTCATGTTTTCACAGCTCTCCAGCAACC 840
QY 841 AAATAGCGCTGTATAGTGTAGACATCTCGGGCTTACGCTTGTCTCTCTAGTG 900
DB 841 AAATAGCGCTGTATAGTGTAGACATCTCGGGCTTACGCTTGTCTCTCTCTAGTG 900
QY 901 TTTCTTATCAGATPACTGCGGAAAGCTTCAATTTTACAAGCCGCAACAGCTTCT 960
DB 901 TTTCTTATCAGATPACTGCGGAAAGCTTCAATTTTACAAGCCGCAACAGCTTCT 960
QY 961 TTGCTAGTGAATTAATGTGTGTGTGTTCCTGATTAAGCAAAATTAATTTAAAAAATG 1020
DB 961 TTGCTAGTGAATTAATGTGTGTGTGTTCCTGATTAAGCAAAATTAATTTAAAAAATG 1020

```

QY 1021 AAAAGTT 1027
DB 1021 AAAAGTT 1027

```

RESULT 5
ADL06443
ID ADL06443 standard; cDNA; 1395 BP.
XX
AC ADL06443;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #23.
XX
KW Human; tumour-associated antigenic target; TAT; cell death; tumour;
XX cancer; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN W02004016225-A2.
XX
PD 26-FEB-2004.
XX
PE 19-AUG-2003; 2003WO-US025892.
XX
PR 19-AUG-2002; 2002US-0404809P.
XX
PR 21-AUG-2002; 2002US-0405645P.
XX
PR 23-SEP-2002; 2002US-0413192P.
XX
PR 15-OCT-2002; 2002US-0419088P.
XX
PR 15-NOV-2002; 2002US-0426847P.
XX
PR 02-JUL-2003; 2003US-0484959P.
XX
PA (GENTECH ) GENENTECH INC.
XX
PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
XX Spencer SD, Wu TD, Zhang Z;
XX
XX WPI; 2004-257144/24.
XX
XX DR P-PSDB; ADL06523.
XX
XX
XX PT New antibody that binds to a tumor-associated antigenic target (TAT)
XX peptide, useful for preparing a composition for diagnosing or
XX treating cancer.
XX
XX PS Claim 1; SEQ ID NO 23; 319pp; English.
XX
XX
XX CC The present invention relates to the isolation of human tumour-associated
XX antigenic target (TAT) polynucleotide and polypeptide sequences. Also
XX CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
XX CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
XX CC humanised antibody. It is conjugated to a growth inhibitory agent. It is
XX CC produced in bacteria or in CHO cells and induces death of a cell to which
XX CC it binds. The antibody is useful for preparing a composition for
XX CC diagnosing or treating tumours and cancer. The present sequence
XX CC represents a human TAT cDNA sequence of the invention.
XX
SQ Sequence 1395 BP; 405 A; 324 C; 286 G; 380 T; 0 U; 0 Other;
Query Match 98.7%; Score 1013.8; DB 12; Length 1395;
Best Local Similarity 99.8%; Pred. No. 3.7e-286;
Matches 1015; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 CAAGAGTTGGGCAAAAATCAAGTATTTGGTCCCGAAACAAAGCTTATCATTAAGAT 63
DB 379 CAAGAGTTGGGCAAAAATCAAGTATTTGGTCCCGAAACAAAGCTTATCATTAAGAT 438
QY 64 AAACAATTATGATGAGTGTGTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAA 123
DB 439 AAACAATTATGATGAGTGTGTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAA 498
QY 124 ACAAAGCTCAAGAGGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTT 183

```

```

Db      ||| 499 ACAAGCTCCAGAAAGCTGGAACATCTTGTCTTCTTGAAATTTTCCCTGAGTT 558
Qy      ||| 184 ATTAAGATACATTTGGCAAGAAAGAGCAACAGATTTCTGGATCCCAAGAGGCAAC 243
Db      ||| 559 ATTAAGATACATTTGGCAAGAAAGAGCAACAGATTTCTGGATCCCAAGAGGCAAC 618
Qy      ||| 244 ACCATGAAGACTAAAGACATACATGAAATTTAGCTGTTAACGGTCCCAAGAAAGTCA 303
Db      ||| 619 ACCATGAAGACTAAAGACATACATGAAATTTAGCTGTTAACGGTCCCAAGAAAGTCA 678
Qy      ||| 304 CTGACAAAGAAACACAGATGATGCTGACATGAGATTAATAAAACGGATGATCA 363
Db      ||| 679 CTGACAAAGAAACACAGATGATGCTGACATGAGATTAATAAAACGGATGATCA 738
Qy      ||| 364 GAAATTAATCTTCTCCATTAAGACGAGATGATCAATGATGCCAAAGACATTTGT 423
Db      ||| 739 GAAATTAATCTTCTCCATTAAGACGAGATGATCAATGATGCCAAAGACATTTGT 798
Qy      ||| 424 TCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTACATG 483
Db      ||| 799 TCAAAAGATGCAAAATGATACACTACTGCTGACAGCTCAAAACACCTCTGCATATTACACG 858
Qy      ||| 484 TACCTCCCTGCTCCTCAAGAGTGGTCTATTGTGCATCATCACTGCTGCTGCTT 543
Db      ||| 859 TACCTCCCTGCTCCTCAAGAGTGGTCTATTGTGCATCATCACTGCTGCTGCTT 918
Qy      ||| 544 AGAAGAACGGCTTTCTGTCGATGAGAGAAATCATTAACAGACGGTGGCAAGAGAGC 603
Db      ||| 919 AGAAGAACGGCTTTCTGTCGATGAGAGAAATCATTAACAGACGGTGGCAAGAGAGC 978
Qy      ||| 604 CATCTTTCTCATCGGTTATTTGCTCCAGAAAGCGTCTTCTGAGAGATCACTGAGGCTTT 663
Db      ||| 979 CATCTTTCTCATCGGTTATTTGCTCCAGAAAGCGTCTTCTGAGAGATCACTGAGGCTTT 1038
Qy      ||| 664 CTCTCTGGGCTTTGGGCAATTTCAAGTTCATGTCATGTCATTTCTCATATATTTGTAATA 723
Db      ||| 1039 CTCTCTGGGCTTTGGGCAATTTCAAGTTCATGTCATGTCATTTCTCATATATTTGTAATA 1098
Qy      ||| 724 CGGTTTCAAAACAGATGGGACACAGAGAACCTCACTCTGTAATTAACAATGAGAAATAGC 783
Db      ||| 1099 CGGTTTCAAAACAGATGGGACACAGAGAACCTCACTCTGTAATTAACAATGAGAAATAGC 1158
Qy      ||| 784 CACGGCGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCCTCCAGCCAAACCCAA 843
Db      ||| 1159 CACGGCGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCCTCCAGCCAAACCCAA 1218
Qy      ||| 844 TAGCGCCGTATAGTGTAGACATCTGCGGCTTCTAGCCCTTCCCTCTTATAGTTC 903
Db      ||| 1219 TAGCGCCGTATAGTGTAGACATCTGCGGCTTCTAGCCCTTCCCTCTTATAGTTC 1278
Qy      ||| 904 TTTAATCAGATTAAGCTGGAAGCTTTCAATTTTACAGCCCTGGAAGCAGCTTTCTTGG 963
Db      ||| 1279 TTTAATCAGATTAAGCTGGAAGCTTTCAATTTTACAGCCCTGGAAGCAGCTTTCTTGG 1338
Qy      ||| 964 CTAGTTGAATTAATGATGATGTTTTCGTAATAAGCAAAATTAATTTAAAAAATG 1020
Db      ||| 1339 CTAGTTGAATTAATGATGATGTTTTCGTAATAAGCAAAATTAATTTAAAAAATG 1395

```

RESULT 6
ADY16296
ID ADY16296 standard; DNA; 1530 BP.

XX AC ADY16296;

XX XX 05-MAY-2005 (first entry)

XX XX DNA encoding a PRO polypeptide, SEQ ID NO 2102.

KW Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antidiabetic; Osteopathic; Hemostatic; Antianemic;
KW Antihypoid; Antidiabetic; Nephroretropic; CNS-Gen.; Hepatotropic;

```

KW Virucide; Gastrointestinal-Gen.; Antipneumonic; Antiallergic; de; gene; diagnosis.
KW Antiallergic; de; gene; diagnosis.
KW Homo sapiens.
XX MO2005016962-A2.
XX
XX 24-FEB-2005.
XX
XX 11-AUG-2004; 2004MO-US026249.
XX
XX 11-AUG-2003; 2003US-0493546P.
XX
XX (GERTH ) GENENTECH INC.
XX
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX WPI; 2005-182330/19.
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
XX treating an immune related disorder, e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
XX Claim 1; SEQ ID NO 2102; 158pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. The polypeptide, agonist or an antagonist, antibody,
XX composition, and method are useful for diagnosing and treating an immune
XX related disorder, e.g. systemic lupus erythematosus, rheumatoid
XX arthritis. The present sequence represents a DNA encoding a PRO
XX polypeptide.
XX
XX Sequence 1530 BP; 435 A; 361 C; 322 G; 412 T; 0 U; 0 Other;
XX
XX Query Match 95.8%; Score 984; DB 14; Length 1530;
XX Best Local Similarity 98.0%; Pred. No. 2; 1e-277;
XX Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy      ||| 12 GGGCAAAAATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAAT 71
Db      ||| 515 GGGAAATTAAGAAAGCTTTGGCAGTGGAAACAACCTGTGTACAGATTAACAACAT 574
Qy      ||| 72 TGATGCAAGATTTTCCCAAGCCCACTATTTTCTTCTTCAATTTGCTGAAACAAAGCT 131
Db      ||| 575 TGATGCAAGATTTTCCCAAGCCCACTATTTTCTTCTTCAATTTGCTGAAACAAAGCT 634
Qy      ||| 132 CCAGAAGGCTGGAACATACCTTTGCTTCTTGAAGAAATTTTCCCTGATGTTAAGAT 191
Db      ||| 635 CCAGAAGGCTGGAACATACCTTTGCTTCTTGAAGAAATTTTCCCTGATGTTAAGAT 694
Qy      ||| 192 ACATTGGCAAGAAAAGAGCAACAGATTTCTGGATCCAGAGAGGGAACACCATGAA 251
Db      ||| 695 ACATTGGCAAGAAAAGAGCAACAGATTTCTGGATCCAGAGAGGGAACACCATGAA 754
Qy      ||| 252 GACTTAAGACATACATGAAATTTAGCTGTTAACGTCGCAAAAAGTACCTGAGCAA 311
Db      ||| 755 GACTTAAGACATACATGAAATTTAGCTGTTAACGTCGCAAAAAGTACCTGAGCAA 814
Qy      ||| 312 AGAACAAGATGTAATGTCGACATGAGATTAATAAAGAGAGTGTGATCAAGAAATTAAT 371
Db      ||| 815 AGAACAAGATGTAATGTCGACATGAGATTAATAAAGAGAGTGTGATCAAGAAATTAAT 874
Qy      ||| 372 CTCTCTCCATTAAGACGAGATGTCATCAAGATGATCCCAAGACATTTGTTAAAAAGA 431
Db      ||| 875 CTCTCTCCATTAAGACGAGATGTCATCAAGATGATCCCAAGACATTTGTTAAAAAGA 934
Qy      ||| 432 TGCAAAATGATCACTACTGCTGACGCTCAAAACACCTCTGCAATTTACATGATCCTCT 491
Db      ||| 935 TGCAAAATGATCACTACTGCTGACGCTCAAAACACCTCTGCAATTTACATGATCCTCT 994
Qy      ||| 492 CCTGCTCTCAAGAGTGTGCTATTTTGGCATCATCAACCTGCTGCTTAGAAGAG 551
Db      ||| 995 CCTGCTCTCAAGAGTGTGCTATTTTGGCATCATCAACCTGCTGCTTAGAAGAG 1054

```

QY 552 GGGTTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTT 611
 CC endotheilium, dermal microvascular endotheilium, pulmonary artery
 Db 1055 GGGTTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTT 1114
 CC endotheilium, myometrium microvascular endotheilium, keratinocyte
 QY 612 CCTCATCGGTTATTTGCTCCTAGAGACGCTCTTCTGAGATCTAGTTGGGCTTTCTTTCTGG 671
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 Db 1115 CCTCATCGGTTATTTGCTCCTAGAGACGCTCTTCTGAGATCTAGTTGGGCTTTCTTTCTGG 1174
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 QY 672 GTTTGGGCAATTTCAATTTCTCATGTGTGTAATTTCTATCTATTTGTATTAACGGTTTTC 731
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 Db 1175 GTTTGGGCAATTTCAATTTCTCATGTGTGTAATTTCTATCTATTTGTATTAACGGTTTTC 1234
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 QY 732 AAACCGATGGGACACAGAGAACCTCACTCTGTATTAACATGAGAAATAGCCAGGCGA 791
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 Db 1235 AAACCGATGGGACACAGAGAACCTCACTCTGTATTAACATGAGAAATAGCCAGGCGA 1294
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 QY 792 TCTCCAGCACCACATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAAACCAATAGCCCT 851
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 Db 1295 TCTCCAGCACCACATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAAACCAATAGCCCT 1354
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 QY 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCCTTCTCTCTTCTTCTTCTTAAATCA 911
 CC for determining the level of RNA expression for a sample, determining the
 Db 1355 GCTATAGTGAACATCTCTGCGGCTTCTAGCCTTCTCTCTTCTTCTTCTTAAATCA 1414
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 QY 912 GATTAACGCTCGGAAGCCTTTCATTTTACAGCCCTGAAGAGCTCTTCTGAGTTGA 971
 CC expression profile is useful in identifying disease pathologies involving
 Db 1415 GATTAACGCTCGGAAGCCTTTCATTTTACAGCCCTGAAGAGCTCTTCTGAGTTGA 1474
 CC alterations of gene expression. The assessment of expression profiles may
 QY 972 ATTATGTGTGTGTTTTCCGTAAATTAAGCAAAATTAATTAATTAATTAATTAATTAATTA 1027
 CC provide meaningful information with respect to tumor type and stage.
 Db 1475 ATTATGTGTGTGTTTTCCGTAAATTAAGCAAAATTAATTAATTAATTAATTAATTAATTA 1530
 CC treatment methods, and prognosis. The gene or protein expression profile
 CC may also be used for creating microarrays. The microarray is useful for
 CC genetic and physical mapping of genomes, DNA sequencing, genetic or
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue
 CC identifications and in identifying promising antibiotics, antiviral or
 CC antifungal agents

RESULT 7
 ID AB235381
 AB235381 standard; cDNA; 1799 BP.
 AC AB235381;
 XX
 DT 05-FEB-2003 (first entry)
 DE Human gene expression profile polynucleotide SEQ ID NO 492.
 XX
 XX Human; artery; endotheilium; umbilical; vein; aorta; pulmonary artery;
 KM bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KM tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KM gene expression; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200274979-A2.
 PD 26-SEP-2002.
 XX
 PF 20-MAR-2002; 2002WO-US008456.
 XX
 PR 20-MAR-2001; 2001US-0276947P.
 XX
 PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 XX
 PI Wan J, Wang Y;
 XX
 DR WPI; 2002-740862/80.
 XX
 XX New gene expression profile generated from primary, endotheilial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer.
 XX
 PS Example 3; Page 634-635; 850bp; English.
 CC
 CC The invention relates to a gene expression profile comprising one or more
 CC genes (AB234809-AB235692) and generated from a cell type. The cell type

CC is a coronary artery endotheilium, umbilical artery or vein endotheilium,
 CC aortic endotheilium, dermal microvascular endotheilium, pulmonary artery
 CC endotheilium, myometrium microvascular endotheilium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies involving
 CC alterations of gene expression. The assessment of expression profiles may
 CC provide meaningful information with respect to tumor type and stage.
 CC treatment methods, and prognosis. The gene or protein expression profile
 CC may also be used for creating microarrays. The microarray is useful for
 CC genetic and physical mapping of genomes, DNA sequencing, genetic or
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue
 CC identifications and in identifying promising antibiotics, antiviral or
 CC antifungal agents

Sequence 1799 BP; 499 A; 434 C; 396 G; 470 T; 0 U; 0 Other;

Query Match 94.3%; Score 968.2; DB 6; Length 1799;
 Best Local Similarity 98.1%; Pred. No. 9.7e-273;
 Matches 990; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 20 AAATCAAGTATTTGCTCCGGAACAAGCTTATCATTAACAGTAAACAATTGATGACAG 79
 Db 791 ATTAAGAACTCTTTGGCAAGTGAACAACCTTGTGTACAGATTAACAACCTTGATGACAG 850
 QY 80 ATGTTTCCCGCAAGCCCACTATTTTCTCTCTCAATGCTGAAACAAAGCTCCAGAAG 139
 Db 851 ATGTTTCCCGCAAGCCCACTATTTTCTCTCTCAATGCTGAAACAAAGCTCCAGAAG 910
 QY 140 CTGGAACATACCTTTGCTCTTGAAGAAATTTTCCCTGATGTTTATTAATGATCATGGC 199
 Db 911 CTGGAACATACCTTTGCTCTTGAAGAAATTTTCCCTGATGTTTATTAATGATCATGGC 970
 QY 200 AAGAAAAGAGAACCAACAGATCT- GGGATCCAGAGGGGAACAACATGAAGACTTAC 258
 Db 971 AAGAAAAGAGAACCAACAGATCTGGGGATCCAGAGGGGAACAACATGAAGACTTAC 1030
 QY 259 GACACATACATGAATTTAGCTGTTAACCGTCCAGAAAAGTCACTGGACAAGAACAC 318
 Db 1031 GACACATACATGAATTTAGCTGTTAACCGTCCAGAAAAGTCACTGGACAAGAACAC 1090
 QY 319 AGATGATATGCTCAGACATGGAATTAATTAACGAGGTGATCAAGAAATTAATCTTCCCT 378
 Db 1091 AGATGATATGCTCAGACATGGAATTAATTAACGAGGTGATCAAGAAATTAATCTTCCCT 1150
 QY 379 CCAATTAAGAGGAGTGCATCAATGATGCCAAGACAAATTTGTTCAAAAGATGCAAT 438
 Db 1151 CCAATTAAGAGGAGTGCATCAATGATGCCAAGACAAATTTGTTCAAAAGATGCAAT 1210
 QY 439 GATACACTAGCTGTCAGCTCACAACACCTGTGCATTTATCATGTATCTCTCTGCTC 498
 Db 1211 GATACACTAGCTGTCAGCTCACAACACCTGTGCATTTATCATGTATCTCTCTGCTC 1270
 QY 499 CTCAAGAGTGTGCTATTTTGGCATCATCACTGCTGTCTGTTGAAGAACGGCTTTC 558
 Db 1271 CTCAAGAGTGTGCTATTTTGGCATCATCACTGCTGTCTGTTGAAGAACGGCTTTC 1330
 QY 559 TGTGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTCTCATC 618
 Db 1331 TGTGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTCTCATC 1390
 QY 619 GGTATTTGCTCTAGAGAGCTCTTCTGAGATCTAGTTGGGCTTTCTTCTGGGTTTGGG 678
 Db 1391 GGTATTTGCTCTAGAGAGCTCTTCTGAGATCTAGTTGGGCTTTCTTCTGGGTTTGGG 1450

QY 679 CCATTTTCAGTTCTCATGTCGTACTATTCTATCATTTATTTGTAATACGGTTTTCACACG 738
DB 1451 CCATTTTCAGTTCTCATGTCGTACTATTCTATCATTTATTTGTAATACGGTTTTCACACG 1510
QY 739 TGGGACACAGAGAACTCAGCTGTATTAACAATGAGAAATAGCAGGAGATCTCAG 798
DB 1511 TGGGACACAGAGAACTCAGCTGTATTAACAATGAGAAATAGCAGGAGATCTCAG 1570
QY 799 CACCAATCTCTCATGTTTTCACAGCTCTCCAGCCCAACCAATATAGCGCTGTATAG 858
DB 1571 CACCAATCTCTCATGTTTTCACAGCTCTCCAGCCCAACCAATATAGCGCTGTATAG 1630
QY 859 TGTAGACATCTCTGCGGCTTTAGCTTGTCCCTCTCTTAGTGTCTTTAATCAGTAACT 918
DB 1631 TGTAGACATCTCTGCGGCTTTAGCTTGTCCCTCTCTTAGTGTCTTTAATCAGTAACT 1690
QY 919 GCCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGCTTCTTGTCTAGTGAATATAGT 978
DB 1691 GCCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGCTTCTTGTCTAGTGAATATAGT 1750
QY 979 GGTGTGTTTTCCGTATATAGCAAAATTAATTTAAAAATGAAAAAGTT 1027
DB 1751 GGTGTGTTTTCCGTATATAGCAAAATTAATTTAAAAATGAAAAAGTT 1799

RESULT 8

AAD38828
ID AAD38828 standard; cDNA; 1418 BP.

XX AC AAD38828;

XX DT 23-SEP-2002 (first entry)

XX DE Human PSNA cDNA #8.

XX KM Human; prostate specific protein; PSP; prostate specific nucleic acid;
KM vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;
KM PSNA; 88.

XX OS Homo sapiens.

XX PN WO200224718-A1.

XX PD 28-MAR-2002.

XX PF 19-SEP-2001; 2001WO-US029386.

XX PR 19-SEP-2000; 2000US-0233746P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Sun Y, Recipon H, Caferkey R, Ali S;

XX DR WPI; 2002-471216/50.

XX PT Novel isolated prostate specific polypeptide useful for identifying,
PT diagnosing, monitoring, staging, imaging, and treating prostate cancer
PT and non-cancerous disease states in prostate.

XX PS Claim 1; Page 187; 210pp; English.

XX CC The invention relates to prostate specific proteins (PSP) and prostate
CC specific nucleic acids (PSNA). Sequences of the invention are useful for
CC identifying, diagnosing, monitoring, staging, imaging and treating
CC prostate cancer and non-cancerous disease states in prostate. They are
CC also useful for producing engineered prostate tissue for treatment and
CC research. The PSNA sequences are used in gene therapy and for producing
CC transgenic animals and cells. The invention is also used as vaccines. The
CC present sequence is human PSNA cDNA

XX SQ Sequence 1418 BP; 418 A; 335 C; 285 G; 380 T; 0 U; 0 Other;

Query Match 92.1%; Score 945.6; DB 6; Length 1418;

Best Local Similarity 97.8%; Pred. No. 3,6e-266;
Matches 1001; Conservative 0; Mismatches 19; Indels 4; Gaps 4;

QY 4 CAAGAGTTGGGCAAAAAATCAAGTATTTGTCCCGAAACAAGTTATCATACAGAT 63
DB 379 CAAGAGTTGGGCAAAAAATCAAGTATTTGTCCCGAAACAAGTTATCATACAGAT 438
QY 64 AAACAATTGATGAGATGTTTCCCGAAGCCCAATATTTCTCTCTCAATGCTGAA 123
DB 439 AAACAATTGATGAGATGTTTCCCGAAGCCCAATATTTCTCTCTCAATGCTGAA 498
QY 124 ACAAGCTCCAGAAAGCTGGAACAATCTTGTCTGTGAAATTTTCCGATGTT 183
DB 499 ACAAGCTCCAGAAAGCTGGAACAATCTTGTCTGTGAAATTTTCCGATGTT 558
QY 184 ATTAAGATACATTGGCAAGAAAGAAAGCAACAAGATTCGAGATCCAGAGGGGAA 243
DB 559 ATTAAGATACATTGGCAAGAAAGAAAGCAACAAGATTCGAGATCCAGAGGGGAA 618
QY 244 ACATGAAAGCTAAAGCAACATACATGAATTTAGCTGTTAACGGTGCAGAAAAGTCA 303
DB 619 ACATGAAAGCTAAAGCAACATACATGAATTTAGCTGTTAACGGTGCAGAAAAGTCA 678
QY 304 CTGACAAAGAAACAAGATGTAATGTCAGACATGAGAAATTAATAACGAGTTGATCAA 363
DB 679 CTGACAAAGAAACAAGATGTAATGTCAGACATGAGAAATTAATAACGAGTTGATCAA 738
QY 364 GAAATTTCTTCTCTCAATTAAGACGAGATGTCATCAAAATGATCCCAAGAAATTTGT 423
DB 739 GAAATTTCTTCTCTCAATTAAGACGAGATGTCATCAAAATGATCCCAAGAAATTTGT 798
QY 424 TCAAAAGATGCAAAATGATACATCTGTCGAGCTCAAAAACCTCTGCAATTAATG 483
DB 799 TCAAAAGATGCAAAATGATACATCTGTCGAGCTCAAAAACCTCTGCAATTAATG 858
QY 484 TACCTCTCTCTCTCTCAAGAGTGTGCTAATTTGCAATCACTGCTGTCTGTT 543
DB 859 TACCTCTCTCTCTCTCAAGAGTGTGCTAATTTGCAATCACTGCTGTCTGTT 918
QY 544 AGAAGAAAGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGTGGCAAGAGAGC 603
DB 919 AGAAGAAAGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGTGGCAAGAGAGC 978
QY 604 CATCTTTCTCTC-ATGCGTATTTGTCCTAGAACGTTCTTGAAGATTAATGTTG 662
DB 979 CATCTTTCTCTC-ATGCGTATTTGTCCTAGAACGTTCTTGAAGATTAATGTTG 1038
QY 663 TCTTTCTGGGTTGGGCAATTCAGTTCTCATGTCGTACTAT-CTATCATTAATGTA- 720
DB 1039 TCTTTCTGGGTTGGGCAATTCAGTTCTCATGTCGTACTAT-CTATCATTAATGTA- 1098
QY 721 TAAAGTTTCAAAACAGTGGGCAACAGAGAACTCACTGTATTAACAATGAGAT 780
DB 1099 TAAAGTTTCAAAACAGTGGGCAACAGAGAACTCACTGTATTAACAATGAGAT 1158
QY 781 AGCCACGCGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAA 840
DB 1159 AGCCACGCGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAA 1218
QY 841 AAATAGGCGCTGCTATGATGTAAGCATCTCGGCTCTAGCTTGTCCCTCTGATG 900
DB 1219 AAATAGGCGCTGCTATGATGTAAGCATCTCGC-CCCTCTAGCTTGTCCCTCTGATG 1277
QY 901 TTTCTTATCAGATACAGCTGGAAGCTTTCAATTTTACAGCCCTGGAAGAGCTCTCT 960
DB 1278 TTTCTTATCAGATACAGCTGGAAGCTTTCAATTTTACAGCCCTGGAAGAGCTCTCT 1337
QY 961 TTGCTAGTGAATATGATGATGTTTTCGATATAGCAAAATTAATTTAAAAATG 1020
DB 1338 CTGCTAGTGAATATGATGATGTTTTCGATATAGCAAAATTAATTTAAAAAT 1397
QY 1021 AAAA 1024
|||

PI	Ma Y, Wang D, Chen R, Xu C, Boyle BJ,	
DR	WPI: 2003-569235/53.	
XX	P-PSDB; ADE07898.	
XX	New polynucleotides, useful for expressing recombinant proteins for	
PT	analysis, characterization or therapeutic use, or as markers for tissues	
PT	in which the corresponding protein is preferentially expressed.	
XX	Claim 1; SEQ ID NO 53; 1177p; English.	
PS		
CC	The invention comprises the amino acid and coding sequences of novel	
CC	proteins. The DNA and protein sequences of the invention are useful as:	
CC	markers for tissues in which the corresponding protein is preferentially	
CC	expressed; as molecular weight markers on gels; as chromosome markers or	
CC	tags; to identify chromosomes or to map related gene positions; and to	
CC	compare with endogenous DNA sequences in patients to identify potential	
CC	genetic disorders. The present DNA sequence represents a gene of the	
CC	invention.	
XX		
XX	Sequence 1560 BP; 444 A; 363 C; 322 G; 431 T; 0 U; 0 Other;	
SO		
Query Match	88.2%; Score 905.6; DB 10; Length 1560;	
Best Local Similarity	91.9%; Pred. NO. 2e-254;	
Matches	987; Conservative 0; Mismatches 39; Indels 48; Gaps 1;	
OY	2 GGCAGAGCTGGGCAAAAAATCAAGTATTTGGTCCCGGAACAAGCTTATCATTCACAG 61	
DB	487 GGACGATAGAGATCTATTATTAAGAACTCTTTGGCAGTGGAAACAACCTGTGTGACAG 546	
OY	62 ATAAACAATTGATGACAGATGTTTCCCGAAGCCACATATTTTCTCTCATTTGCTG 121	
DB	547 ATAAACAATTGATGACAGATGTTTCCCGAAGCCACATATTTTCTCTCATTTGCTG 606	
OY	122 AAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTGGAGAAATTTTCCCTGATG 181	
DB	607 AAACAAAGCTCCAGAAAGCTGGAACATACCTTTGCTTGGAGAAATTTTCCACATTA 666	
OY	182 TTATTAGATACATTGGCAAGAAAAAGAGCAACAGATTTCTGGATCCGAGAGGGA 241	
DB	667 TTATTAGATACATTGGCAAGAAAAAGAGCAACAGATTTCTGGATCCGAGAGGGA 726	
OY	242 ACACCATGAAGCTTACGACACATACATGAATTTAGCTGGTTAACGGTCCAGAAAGT 301	
DB	727 ACACCATGAAGCTTACGACACATACATGAATTTAGCTGGTTAACGGTCCAGAAAGT 786	
OY	302 CACTGGAACAAGAACACAGATGTAATCGTCAGACATGAATAATTAACGAGTGCATC 361	
DB	787 CACTGGAACAAGAACACAGATGTAATCGTCAGACATGAATAATTAACGAGATGATC 846	
OY	362 AAGAATTTATCTTCTCTCCATTAAGAC----- 389	
DB	847 AAGAATTTATCTTCTCTCCATTAAGACATGTCACACAGTGGATCCCAAGACACTT 906	
OY	390 -----GATGTCATCAATGATCCCAAGACACTTGTCAAAAGATG 433	
DB	907 ATTCAAAAGATGCAATGATGTCATCAAAAGATCCCAAGACACTTGTCAAAAGATG 966	
OY	434 CAAATGATACCTACTGCTGACGCTCAAAACACCTCTGCATATTAACATGTAACCTCC 493	
DB	967 CAAATGATACCTACTGCTGACGCTCAAAACACCTCTGCATATTAACATGTAACCTCC 1028	
OY	494 TGCTCTCAAGAGTGTGGTCTATTTTGGCATCATCACTGCTGTCTGTTAGAAAGCG 553	
DB	1027 TGCTCTCAAGAGTGTGGTCTATTTTGGCATCATCACTGCTGTCTGTTAGAAAGCG 1086	
OY	554 CTTTCTGCTCAATGAGAGAAATCATTAACAGCGTGGCAACAAGAGGCAATCTTTTCC 613	
DB	1087 CTTTCTGCTCAATGAGAGAAATCATTAACAGCGTGGCAACAAGAGGCAATCTTTTCC 1146	
OY	614 TCATCGGTTATTTGCTCCCTTAAGAAAGCTCTTGAAGATCTAATGTGGCTTTCTTCTGG 673	
DB	1147 TCATCGGTTATTTGCTCCCTTAAGAAAGCTCTTGAAGATCTAATGTGGCTTTCTTCTGG 1206	

Oy		674	TTGGGCATTTCCAGTTCCTCANGTGTAATTCTATCATTAATGTAACGGTTTTCAA	733
Dd		1207	TTGGGCCAATTCAGTTCCTCACTGTGTACTATTTCTATCATTAATGTTTTCAA	1286
Oy		734	ACCAGTGGGACACACAGAAACCTCACTCTGTAAATPAACAATGAGGAATAGCCCGCATC	793
Dd		1267	ACCAAGTGGGGACACAGAAACCTCACTCTGTAAATPAACAATGAGGAATAGCCATGCGATC	1326
Oy		794	TCCAGCACCAATCTCTCCATGTTTTCCACAGCTCTCTCAGCCAAACCCAATATGGGCTGC	853
Dd		1327	TCCAGCACCAATCTCTCCATGTTTTCCACAGCTCTCTCAGCCAAACCCAATATGGGCTGC	1386
Oy		854	TATAGTGTAGAATCCTCGCGGCTTCTAGCCTGTGCTCCTCTTAGTGTCTTTAATCAGA	913
Dd		1387	TATAGTGTAGAATCCTCGCGGCTTCTAGCCTGTGCTCCTCTTAGTGTCTTTAATCAGA	1446
Oy		914	TAACTGCTCTGAAGCCTTTTCATTTTACAGCCCTGAAGAGTCTTCTTGTCTAGTGAAT	973
Dd		1447	TAACTGCTCTGAAGCCTTTTCATTTTACAGCCCTGAAGAGTCTTCTTGTCTAGTGAAT	1506
Oy		974	TATGTGTGTGTTTTCCTGTAATPAGAAAATAATTAAAAAATGAAGATT	1027
Dd		1507	TATGTGTGTGTTTTCCTGTAATPAGAAAATAATTAAAAAATGAAGATT	1560
RESULT 12				
ABL65007				
ID	ABL65007	standard; DNA; 1421 BP.		
AC				
XX	ABL65007;			
DT	15-MAY-2002	(first entry)		
DE	Lung cancer related gene sequence SEQ ID NO:3144.			
XX				
KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;			
KV	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;			
KM	cyclostatic; gene therapy; anti-neoplastic; Wilm's tumour; adenocarcinoma;			
KX	gene; ds.			
OS	Homo sapiens.			
PN	MO200194629-A2.			
PD	13-DEC-2001.			
XX				
PF	30-MAY-2001; 2001WO-USO10838.			
XX				
PR	05-JUN-2000; 2000US-0209473P.			
PR	05-JUN-2000; 2000US-0209531P.			
PR	18-SEP-2000; 2000US-0233133P.			
PR	18-SEP-2000; 2000US-0233617P.			
PR	20-SEP-2000; 2000US-0234009P.			
PR	20-SEP-2000; 2000US-0234034P.			
PR	20-SEP-2000; 2000US-0234052P.			
PR	22-SEP-2000; 2000US-0234509P.			
PR	22-SEP-2000; 2000US-0234567P.			
PR	25-SEP-2000; 2000US-0234923P.			
PR	25-SEP-2000; 2000US-0234924P.			
PR	25-SEP-2000; 2000US-0235077P.			
PR	25-SEP-2000; 2000US-0235082P.			
PR	25-SEP-2000; 2000US-0235134P.			
PR	25-SEP-2000; 2000US-0235280P.			
PR	26-SEP-2000; 2000US-0235637P.			
PR	26-SEP-2000; 2000US-0235638P.			
PR	27-SEP-2000; 2000US-0235711P.			
PR	27-SEP-2000; 2000US-0235720P.			
PR	27-SEP-2000; 2000US-0235840P.			
PR	27-SEP-2000; 2000US-0235863P.			
PR	28-SEP-2000; 2000US-0236028P.			
PR	28-SEP-2000; 2000US-0236032P.			
PR	28-SEP-2000; 2000US-0236033P.			

PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237596P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppe DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 PS
 PS Claim 1; SEQ ID NO 3344; 44bp; English.
 XX
 XX The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
 CC to AB170110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytotoxic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, esophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
 CC tumour
 CC
 XX Sequence 1421 BP; 419 A; 325 C; 292 G; 385 T; 0 U; 0 Other;
 SO
 Query Match 87.8%; Score 902.2; DB 6; Length 1421;
 Best Local Similarity 92.8%; Pred. No. 1.9e-253;
 Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

Db 609 ACACTATCATGAATTTAGCTGTAAAGGTGCAGAAAGTCACTGGACAAAGAACACA 668
 Qy 320 GATGTATCGTCAGACATGAGATTAATTAACGAGAGTTGATCAAGAAATTTATCTTCTC 379
 Db 669 GATGTATCGTCAGACATGAGATTAATTAACGAGAGTTGATCAAGAAATTTATCTTCTC 728
 Qy 380 CAATTAAGAC-----GG 391
 Db 729 CAATTAAGACAGATGTCACCAAGTGGATCCCAAGACAGTTATTCAAGATGCAATG 788
 Qy 392 ATGTATCACAATGATATCCCAAGACATTTGTTCAAAAGATGCAATGATCTACTGC 451
 Db 789 ATGTATCACAATGATATCCCAAGACATTTGTTCAAAAGATGCAATGATCTACTGC 848
 Qy 452 TGCAGCTCAAAACACCTCTGCATATTAATACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 511
 Db 849 TGCAGCTCAAAACACCTCTGCATATTAATACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 908
 Qy 512 TCTATTTTGGCATCATCATCCTGCTGCTGCTTGAAGACGGCTTTCTGCTGCATGAG 571
 Db 909 TCTATTTTGGCATCATCATCCTGCTGCTGCTTGAAGACGGCTTTCTGCTGCATGAG 968
 Qy 572 AGAATATATAACAGACGGTGGCAACAGAGGCATTTTCTCATGCGTTATGCTCCT 631
 Db 969 AGAATATATAACAGACGGTGGCAACAGAGGCATTTTCTCATGCGTTATGCTCCT 1028
 Qy 632 AGAAGCTCTTTCAGAGATCTAGTTGGCTTTCTTCTGAGGTTGGCCATTTCACTTCT 691
 Db 1029 AGAAGCTCTTTCAGAGATCTAGTTGGCTTTCTTCTGAGGTTGGCCATTTCACTTCT 1088
 Qy 692 CATGTGTACTATTTCTATCATTTATTTGATPAACGGTTTCAACCAAGTGGGCACACAG 751
 Db 1089 CATGTGTACTATTTCTATCATTTATTTGATPAACGGTTTCAACCAAGTGGGCACACAG 1148
 Qy 752 AACCTACTCTGTAATTAACATAGAGAAATGCCAGCGCATTTCCACACCAATCTCTCC 811
 Db 1149 AACCTACTCTGTAATTAACATAGAGAAATGCCAGCGCATTTCCACACCAATCTCTCC 1208
 Qy 812 ATGTTTCCACAGCTCTCCAGCAACCAACCAATAGCCCTGCTATGATGATGATCTG 871
 Db 1209 ATGTTTCCACAGCTCTCTCCAGCAACCAACCAATAGCCCTGCTATGATGATGATCTG 1268
 Qy 872 CGGCTTACGCTTGTCT 931
 Db 1269 CGGCTTACGCTTGTCT 1328
 Qy 932 TCAATTTTACAGCGCTTGAAGCAGCTCTTCTGTAAGTTGAATATGAGTGTGTTTCC 991
 Db 1329 TCAATTTTACAGCGCTTGAAGCAGCTCTTCTGTAAGTTGAATATGAGTGTGTTTCC 1388
 Qy 992 GTAATAGCAAAATTAATTTAAAAAATGAAA 1024
 Db 1389 GTAATAGCAAAATTAATTTAAAAAATGAAA 1421
 RESULT 13
 ADR65880
 ID ADR65880 standard; DNA; 1421 BP.
 XX ADR65880,
 XX 02-DEC-2004 (first entry)
 XX
 XX Human prostatic carcinoma derived DNA SEQ ID 76 #1.
 XX human, cytostatic; diagnosis; prostatic cancer;
 XX differential expression analysis; de.
 XX Homo sapiens.
 XX W02004076614-A2.
 XX

PD 10-SEP-2004.
 XX 22-FEB-2004; 2004WO-DE000433.
 XX 27-FEB-2003; 2003DE-01009985.
 PR 14-MAY-2003; 2003DE-01022134.
 XX (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (ROSE/) ROSENTHAL A.
 PA (HERM/) HERMANN K.
 PA (PILA/) PILARSKY C.
 XX
 PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarzky C, Specht T;
 PI Schmitt A, Beckmann G, Bruemendorf T, Kimmernann H, Roeppcke S;
 PI Kinzhong U, Staub E;
 XX WPI; 2004-653386/63.
 DR
 XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 PT agents.
 XX
 XX Claim 1; Page 282; 1607bp; German.
 PS
 XX This invention describes novel cytosolic polynucleotide and polypeptide
 CC sequences which can be used in a method for diagnosing prostatic cancer
 CC or the risk of developing prostatic cancer. Diagnosis is based on
 CC determining over transcription or over expression of the sequences in
 CC prostatic tissue. Screening for inhibitors of the sequences or detection
 CC substances involves a binding assay, any compounds that bind are
 CC selected, optionally after deconvolution of mixtures. Detection of a
 CC predetermined minimum level of the reporter indicates the presence of
 CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
 CC short-interfering RNA or ribozymes; an organic molecule of molecular
 CC weight below 5000, preferably 300, that binds to the polypeptide; an
 CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
 CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
 CC (monoclonal) antibody directed against Ab or any of the above derivatised
 CC with a reporter group, cell toxin, immunostimulatory molecules and/or
 CC radioisotope. The polynucleotides are identified in human prostatic
 CC cancer by differential expression analysis, using DNA microarrays,
 CC between normal and tumorous tissues, with (over)expression being detected
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that
 CC CD4 was upregulated in many of them. Sections of tissue, isolated from
 CC prostatic cancer patients, or subjects at risk, were incubated
 CC sequentially with anti-human CD4 murine monoclonal antibodies;
 CC biotinylated second antibody; streptavidin-conjugated horseradish
 CC peroxidase and then diaminobenzidine as colour former (brown). The
 CC samples were counterstained with hemalum (blue). Malignant cells stained
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
 CC lymph node metastases were also stained. ADR65805-ADR6654 represent the
 CC polynucleotide and polypeptide sequences used in the method of the
 CC invention.
 CC
 XX
 SQ Sequence 1421 BP; 419 A; 325 C; 292 G; 385 T; 0 U; 0 Other;
 Query Match 87.8%; Score 902.2; DB 13; Length 1421;
 Best Local Similarity 92.8%; Pred. No. 1.9e-253;
 Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

QY 200 AAGAAAAGAGAGAACAGATTCTGGGATCCCGAGGGGAAACCACTGAAAGCTAAGC 259
 DB 549 AAGAAAAGAGAGAACAGATTCTGGGATCCCGAGGGGAAACCACTGAAAGCTAAGC 608
 QY 260 ACAATATCATGAAATTTAGCTGGTTAACGGTGCAGAAAAGTCACTGGACAAAGAACACA 319
 DB 609 ACAATATCATGAAATTTAGCTGGTTAACGGTGCAGAAAAGTCACTGGACAAAGAACACA 668
 QY 320 GATGATCGTCACAGCATGAGAAATATATATAACGAGTGTATCAAGAAATATCTTCTC 379
 DB 669 GATGATCGTCACAGCATGAGAAATATATATAACGAGTGTATCAAGAAATATCTTCTC 728
 QY 380 CAATTAAGAC-----GG 391
 DB 729 CAATTAAGACAGATGTCACACAGTGGATCCCAAGACAGTTATTCAAAAGATGCAATG 788
 QY 392 ATGTCATCAATGATATCCCAAGACAAATGTTCAAAAAGATGCAATGATACCTAGTCG 451
 DB 789 ATGTCATCAATGATATCCCAAGACAAATGTTCAAAAAGATGCAATGATACCTAGTCG 848
 QY 452 TGCAGCTCAAAACACTCTGCAATTAATCAATGATCTCTCTGCTCTCAAGAGTGG 511
 DB 849 TGCAGCTCAAAACACTCTGCAATTAATCAATGATCTCTCTGCTCTCAAGAGTGG 908
 QY 512 TCTATTTTGCATCATATCACTGCTGCTGCTTGAAGAAAGGCTTCTGCTGCAATGAG 571
 DB 909 TCTATTTTGCATCATATCACTGCTGCTGCTTGAAGAAAGGCTTCTGCTGCAATGAG 968
 QY 572 AGAATCATATAACAGACGGTGCACAAAGAGGCAATCTTCTCATCGTTATGTCCT 631
 DB 969 AGAATCATATAACAGACGGTGCACAAAGAGGCAATCTTCTCATCGTTATGTCCT 1028
 QY 632 AGAAGCTCTTCTGAGGATCTAGTGGGCTTTCTTCTGCGTTTGGCCATTCAGTTCT 691
 DB 1029 AGAAGCTCTTCTGAGGATCTAGTGGGCTTTCTTCTGCGTTTGGCCATTCAGTTCT 1088
 QY 692 CAGTGTGATCTATTCATATATATGTAATGATGATTAAGGTTTAAACCGATGGGACACAG 751
 DB 1089 CAGTGTGATCTATTCATATATATGTAATGATGATTAAGGTTTAAACCGATGGGACACAG 1148
 QY 752 AACCTCACTGTATTAATCAATGAGAAATAGCAAGCGCATCTCCAGCAACCAATCTCTCC 811
 DB 1149 AACCTCACTGTATTAATCAATGAGAAATAGCAAGCGCATCTCCAGCAACCAATCTCTCC 1208
 QY 812 ATGTTTCCACAGCTCTCCAGCCCAACCAATAGCGCTGTATAGTATGACATCTCG 871
 DB 1209 ATGTTTCCACAGCTCTCCAGCCCAACCAATAGCGCTGTATAGTATGACATCTCG 1268
 QY 872 CGGCTTACCTTGTGCTCTCTCTTATGTTCTTTATTCAGATPACCTCGGAAGCTT 931
 DB 1269 CGGCTTACCTTGTGCTCTCTCTTATGTTCTTTATTCAGATPACCTCGGAAGCTT 1328
 QY 932 TCAATTTACAGCGCCGGAAGCAGCTCTTGTGATGTAATATGATGATGTTTTC 991
 DB 1329 TCAATTTACAGCGCCGGAAGCAGCTCTTGTGATGTAATATGATGATGTTTTC 1388
 QY 992 GTAATAGCAAAATTAATTTAAAAAATGAAAA 1024
 DB 1389 GTAATAGCAAAATTAATTTAAAAAATGAAAA 1421

RESULT 14
 ADR66783
 ID ADR66783 standard; DNA; 1421 BP.
 XX
 AC ADR66783;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human prostatic carcinoma derived DNA SEQ ID 76 #4.
 XX
 KW human, cytostatic; diagnosis; prostatic cancer;

KW differential expression analysis; ds.
XX Homo sapiens.
OS
XX MO2004076614-A2.
PN
XX 10-SEP-2004.
PD
XX 22-FEB-2004; 2004WO-DE000433.
PF
XX 27-FEB-2003; 2003DE-0100985.
PR 14-MAY-2003; 2003DE-01022134.
XX
XX (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T,
PI Schmitt A, Beckmann G, Bruemendorf T, Kinneemann H, Roepcke S,
PI Xinhong L, Staub E;
XX
XX WPI: 2004-653386/63.
XX
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX
XX Claim 1; Page 1452; 1607pp; German.
XX
XX This invention describes novel cytosolic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminebenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membranes and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADB65805-ADB65954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX
XX Sequence 1421 BP; 419 A; 325 C; 292 G; 385 T; 0 U; 0 Other;
SQ
Query Match 87.8%; Score 902.2; DB 13; Length 1421;
Best Local Similarity 92.8%; Pred. No. 1.9e-253;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

Db 429 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATGCTGAAACAACTCCAGAGG 488
Qy 140 CTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATACATTGGC 199
Db 489 CTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCAGATATTTATTAATACATTGGC 548
Qy 200 AAGAAAAGAGACACACGATTTCTGGGATCCAGAGAGGGGAAACCATGAAGACTAACG 259
Db 549 AAGAAAAGAGACACACGATTTCTGGGATCCAGAGAGGGGAAACCATGAAGACTAACG 608
Qy 260 ACACATGATGAAATTTAGCTGTGTTACGGGTGCAGAAAGTCACTGGAACAAAGAACACA 319
Db 609 ACACATGATGAAATTTAGCTGTGTTACGGGTGCAGAAAGTCACTGGAACAAAGAACACA 668
Qy 320 GATGATGCTCAGACATGAGAAATTAATTAATAACGAGTGTGATCAAGAAATTAATCTTCTC 379
Db 669 GATGATGCTCAGACATGAGAAATTAATTAATAACGAAATTAATCAAGAAATTAATCTTCTC 728
Qy 380 CAATTAAGAC-----GG 391
Db 729 CAATTAAGACAGATGTCACCAAGTGGATCCAAAGACAGTTATTCAAAAGATGCAATG 788
Qy 392 ATGTCATCAAAATGGAATCCAAAGACAAATTTGTTCAAAAGATGCAAAATGACTACTGC 451
Db 789 ATGTCATCAAAATGGAATCCAAAGACAAATTTGTTCAAAAGATGCAAAATGACTACTGC 848
Qy 452 TGCAGCTCACAACACACCTCTGATATTATTAATGATCTCTCTGCTCTCAAGAGTGG 511
Db 849 TGCAGCTCACAACACACCTCTGATATTATTAATGATCTCTCTGCTCTCAAGAGTGG 908
Qy 512 TCTATTTTGCATCATCATCTGCTGCTGCTTGAAGACAGGCTTCTGCTGCAATGAG 571
Db 909 TCTATTTTGCATCATCATCTGCTGCTGCTTGAAGACAGGCTTCTGCTGCAATGAG 968
Qy 572 AGAAATATTAACAGCGGTGGCAAGAGGCCATTTTCTCATGTTATTTGCTCT 631
Db 969 AGAAATATTAACAGCGGTGGCAAGAGGCCATTTTCTCATGTTATTTGCTCT 1028
Qy 632 AGAAGCTCTTTCGAGGATCTAGTTGGGCTTTCTTCTGGGTTGGGCCATTTCACTTCT 691
Db 1029 AGAAGCTCTTTCGAGGATCTAGTTGGGCTTTCTTCTGGGTTGGGCCATTTCACTTCT 1088
Qy 692 CATGTGTGATCTATTTGATATTTGATTAACGTTTTCACACAGTGGCACACAGAG 751
Db 1089 CATGTGTGATCTATTTGATATTTGATTAACGTTTTCACACAGTGGCACACAGAG 1148
Qy 752 AACCTCACTCTGTAATTAACAATGGAATAGCCAGCGGAGATCTCCACACCAATCTCTCC 811
Db 1149 AACCTCACTCTGTAATTAACAATGGAATAGCCAGCGGAGATCTCCACACCAATCTCTCC 1208
Qy 812 ATGTTTTCACAGCTCTCCAGCCCAACCAATAGCCCTGCTTAATGATGATCATCTTG 871
Db 1209 ATGTTTTCACAGCTCTCTCCAGCCCAACCAATAGCCCTGCTTAATGATGATCATCTTG 1268
Qy 872 CGGCTTTCAGCTTGTGCTCTTCTTGAATGTTCTTTAATCAAGATAGCTGGAAGCTT 931
Db 1269 CGGCTTTCAGCTTGTGCTCTTCTTGAATGTTCTTTAATCAAGATAGCTGGAAGCTT 1328
Qy 932 TCAATTTTACAGCGCTTGAAGAGCTCTTCTGATGTAATTAATGAGTGTGTTTCC 991
Db 1329 TCAATTTTACAGCGCTTGAAGAGCTCTTCTGATGTAATTAATGAGTGTGTTTCC 1388
Qy 992 GTAATTAAGCAAAATTAATTTTAAATAATGAAAA 1024
Db 1389 GTAATTAAGCAAAATTAATTTTAAATAATGAAAA 1421
RESULT 15
ABK64529
ID ABK64529 standard; DNA; 1586 BP.
XX
AC ABK64529;

XX 18-JUN-2002 (first entry)
DT Human benign prostatic hyperplasia gene #424.
XX Human benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX Homo sapiens.
OS MO200212440-A2.
XX 14-FEB-2002.
PD 07-AUG-2001; 2001WO-US024708.
XX 07-AUG-2000; 2000US-0223323P.
XX 05-JUN-2001; 2001US-00873319.
XX (GENE-) GENE LOGIC INC.
PA (NISR) JAPAN TOBACCO INC.
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
XX cells.
XX Disclosure; Page 247-248; 444pp; English.
XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles. (I)
CC is useful for diagnosing the onset or progression of BPH. (II) is useful
CC for identifying an agent that modulates the onset or progression of BPH.
CC The methods are useful to present information identifying the expression
CC level in a tissue or cells, by comparing the expression level of genes
CC given in the specification in the tissue or cells to the level of
CC expression of gene in the database, and displaying the expression levels
CC of at least one gene in the tissue or cell sample compared to the
CC expression level in BPH. Agents using (II) are useful for treating BPH or
CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic
CC hyperplasia gene sequences of the invention
XX
SQ Sequence 1586 BP; 449 A; 369 C; 330 G; 438 T; 0 U; 0 Other;
Query Match 81.6%; Score 838; DB 6; Length 1586;
Best Local Similarity 88.4%; Pred. No. 1.3e-234;
Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;
QY 20 AAATCAAGTATTTGTCCTCCGAAACAAGCTATCATTAAGATAAACAATGATGAG 79
DB 499 ATAGAAACTTTTGGAGCTGGAACAACACTTTGTCACAGATAAACAATTTATGAG 558
QY 80 ATGTTTCCCAAGCCCACTATTTTCTTCTTCATTTGCTGAACAAGCTCCAGAAG 139
DB 559 ATGTTTCCCAAGCCCACTATTTTCTTCTTCATTTGCTGAACAAGCTCCAGAAG 618
QY 140 CTGAAACATTAATCTTTGCTTTTGAAGAAATTTTCCCTGATTTATTAATTAATGATGAG 199
DB 619 CTGAAACATTAATCTTTGCTTTTGAAGAAATTTTCCAGATATTATTAAGATTAATGAG 678

QY 200 AAGAAAAGAGCAACAGATTCGGGATCCAGAGGGGAAACCATGAGAATAACG 259
DB 679 AAGAAAAGAGCAACAGATTCGGGATCCAGAGGGGAAACCATGAGAATAACG 738
QY 260 ACACATCATGAATAATTAGCTGTGAACGGTCCAGAAAAGTCACTGACAAAGACACA 319
DB 739 ACACATCATGAATAATTAGCTGTGAACGGTCCAGAAAAGTCACTGACAAAGACACA 798
QY 320 GATGTATGTCAGCATGAGAAATTAATAAAGGAGTGTGATCAAGAAATTAATCTTCTC 379
DB 799 GATGTATGTCAGCATGAGAAATTAATAAAGGAGTGTGATCAAGAAATTAATCTTCTC 858
QY 380 CAATTAAGAC----- 389
DB 859 CAATTAAGACAGATGTACCAAGTGTATCCAAAGACAGTTATTAAAAAGATCAATG 918
QY 390 -----GATGTATCAAA 403
DB 919 ATGTACCAAGTGTATCCAAATTAATTAATCAAGATGTCAAAATGATGTATCAAA 978
QY 404 TGGATCCCAAGCAATTTGTTCAAAAAGATGCAATGATCACTACCTGAGCTCAAA 463
DB 979 TGGATCCCAAGCAATTTGTTCAAAAAGATGCAATGATCACTACCTGAGCTCAAA 1038
QY 464 ACACTCTGATATTATCATGATACCTCTCTGCTCTCAAGAGTGTGATATTGSCA 523
DB 1039 ACACTCTGATATTATCATGATACCTCTCTGCTCTCAAGAGTGTGATATTGSCA 1098
QY 524 TCATCACTGCTGTCTGTGAAGAAAGCGCTTCTGCTGATGAGAAATCATPAC 583
DB 1099 TCATCACTGCTGTCTGTGAAGAAAGCGCTTCTGCTGATGAGAAATCATPAC 1158
QY 584 AGACGGTGGCAAGAGGCGATCTTCCATCGGTTATTTGCTCCAGAGGCTCTC 643
DB 1159 AGACGGTGGCAAGAGGCGATCTTCCATCGGTTATTTGCTCCAGAGGCTCTC 1218
QY 644 TGAAGATCTAGTGGGCTTCTTCTGCGTGGGCAATTCAGTTCATGTTGTA 703
DB 1219 TGAAGATCTAGTGGGCTTCTTCTGCGTGGGCAATTCAGTTCATGTTGTA 1278
QY 704 ATTCTATCATTTATTTGTAACGGTTTCAACACAGTGGGCAACAGAAACCTCACTG 763
DB 1279 ATTCTATCATTTATTTGTAACGGTTTCAACACAGTGGGCAACAGAAACCTCACTG 1338
QY 764 TAATTAACAATGAGAAATAGCAAGCGATCTCCAGACCAATCTCTCATGTTTCCACA 823
DB 1339 TAATTAACAATGAGAAATAGCAAGCGATCTCCAGACCAATCTCTCATGTTTCCACA 1398
QY 824 GCTCCTCCAGCAACCAATAGGCGCTGCTATGATAGATCTGCGGCTCTAGCC 883
DB 1399 GCTCCTCCAGCAACCAATAGGCGCTGCTATGATAGATCTGCGGCTCTAGCC 1458
QY 884 TTGTCCTCTCTTAGTGTCTTTAATCAGATTAATGCTGGAAGCCCTTCAATTTACAG 943
DB 1459 TTGTCCTCTCTTAGTGTCTTTAATCAGATTAATGCTGGAAGCCCTTCAATTTACAG 1518
QY 944 CCTGGAAGAGTCTTCTTGTAGTGAATTAATGATGATGTTTCCGTAATAGCAAA 1003
DB 1519 CCTGGAAGAGTCTTCTTGTAGTGAATTAATGATGATGTTTCCGTAATAGCAAA 1578
QY 1004 ATAAATTT 1011
DB 1579 ATAAATTT 1586

Search completed: December 10, 2005, 20:46:12
Job time : 671 secs

THIS PAGE BLANK (USPTO)

1	1027	100.0	1027	6	CO9833344
2	1027	100.0	1027	6	AF151103
3	1025.4	99.8	1074	6	AX074415
4	1010	98.3	1825	6	BC072396
5	1008.8	98.2	1054	6	BC062761
6	984	95.8	1530	6	CS032596
7	984	95.8	1530	6	CS041548
8	984	95.8	1530	6	HM073878
9	973	94.7	1470	6	BC072387
10	938	91.3	958	8	HM073838
11	903.8	88.0	1655	8	BC039725
12	902.2	87.8	1441	6	CO986252
13	902.2	87.8	1441	6	AX332835
14	892.2	87.8	1441	6	HM073838
15	877.4	85.4	1033	8	HM073838
16	865.2	84.2	1402	8	HM073838
17	857.4	83.5	1003	8	HM073838
18	842.4	82.0	1757	8	BC039116

19	838	81.6	1586	6	CS032845
20	838	81.6	1586	6	CS041797
21	838	81.6	1586	8	HDMTCRGAD
22	821.6	80.0	1046	8	HDMTCGGH
23	821.6	80.0	1046	10	G31759
24	795.8	77.5	1162	6	CQ489650
25	795.8	77.5	1162	6	CQ492275
26	795.8	77.5	1162	6	CQ492662
27	795.8	77.5	1162	6	CQ492666
28	795.8	77.5	1162	6	CQ492973
29	795.8	77.5	1162	6	CQ493284
30	795.8	77.5	1162	6	CQ495492
31	795.8	77.5	1162	6	CQ495873
32	733.4	71.4	757	8	HDMTCGGJ
33	619	60.3	924	6	CQ493214
34	601	58.5	766	8	HDMTCRGAC
35	589	57.4	847	8	HDMTCGCD3
36	589	57.4	140631	8	AF159056
37	589	57.4	171816	8	AC006033
38	579.6	56.4	586	8	HDMTCGNO
39	578.2	56.3	1041	8	AY190025
40	577.2	56.2	1041	8	AY190027
41	576.4	56.1	66558	14	AC130306
42	522.8	50.9	1080	6	CS036408
43	522.8	50.9	1080	6	CS045360
44	522.8	50.9	1080	6	HSTRRCR
45	516.4	50.5	1080	6	AR364341
					CS032845 Sequence
					CS041797 Sequence
					M30894 Human T-cell
					M17323 Homo sapien
					G31759 SMS51280 Ex
					CQ489650 Sequence
					CQ492275 Sequence
					CQ492662 Sequence
					CQ492666 Sequence
					CQ492973 Sequence
					CQ493284 Sequence
					CQ495492 Sequence
					CQ495873 Sequence
					M17325 Homo sapien
					CQ493214 Sequence
					M27333 Homo sapien
					M17998 Human T-cell
					AF159056 Homo sapi
					AC006033 Homo sapi
					M15007 Human T-cell
					AY190025 Macaca mu
					AY190027 Macaca mu
					AC130306 Homo sapi
					CS036408 Sequence
					CS045360 Sequence
					Y00790 Human mRNA
					AR364341 Sequence

ALIGNMENTS

LOCUS	LOCUS	LOCUS	LOCUS
CO983344	CO983344	1027 bp	DNA
DEFINITION	Sequence 2 from Patent WO2005000889.		linear
ACCESSION	CO983344		PAT 25-JAN-2005
VERSION	CO983344.1	GI:58191770	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Berzofsky, J. A., Oh, S. and Paetan, I.		
TITLE	Immunogenic peptides for the treatment of prostate and breast cancer		
JOURNAL	Patent: WO 2005000889-A 2 06-JAN-2005;		
	National Institutes of Health (US)		
FEATURES	Location/Qualifiers		
source	1..1027		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	100.0%;	Score 1027;	DB 6;
Best Local Similarity	100.0%;	Prod. No. 3.5e-253;	Length 1027;
Matches 1027;	Conservative	0;	Mismatches 0;
		Indels	0;
		Gaps	0;

ORIGIN

100.0%; Score 1027; DB 6; Length 1027;

```
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Oy 181 GTTATTAAAGATACATTGGCAAGAAAAGACACACGATTCTGGGATCCCGAGAGGG 240
Db 181 GTTATTAAAGATACATTGGCAAGAAAAGACACACGATTCTGGGATCCCGAGAGGG 240
Oy 241 AACCCAGTGAAGACTTAACGACATACATGAAATTTAGCTGGTTAAGGGTGCAGAAAAG 300
Db 241 AACCCAGTGAAGACTTAACGACATACATGAAATTTAGCTGGTTAAGGGTGCAGAAAAG 300
Oy 301 TCACGTGCAAAAGACACAGATGTATCCGTACAGACATGAGAAATTAATAAAAACGAGTTGAT 360
Db 301 TCACGTGCAAAAGACACAGATGTATCCGTACAGACATGAGAAATTAATAAAAACGAGTTGAT 360
Oy 361 CAAGAAATTAATCTTCTCTCCATTAAGACGAGTGTATCAACAATGATCCCAAGACAT 420
Db 361 CAAGAAATTAATCTTCTCTCCATTAAGACGAGTGTATCAACAATGATCCCAAGACAT 420
Oy 421 TGTTCAAAAGATGCAAAATGATACACTAGCTGACGCTCAACAAACACTCTGCATATTAC 480
Db 421 TGTTCAAAAGATGCAAAATGATACACTAGCTGACGCTCAACAAACACTCTGCATATTAC 480
Oy 481 ATGTACCTCTCTGCTGCTCTCAAGAGTGTGTATTTTGGCATCATCAGCTGCTGTCTG 540
Db 481 ATGTACCTCTCTGCTGCTCTCAAGAGTGTGTATTTTGGCATCATCAGCTGCTGTCTG 540
Oy 541 CTTAAGACAGCGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAACAAG 600
Db 541 CTTAAGACAGCGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAACAAG 600
Oy 601 GGCCATCTTTTCCCTCATGGTATTATGTCCTTGAAGAGCGCTTTGAGAGATCTAGTGGGC 660
Db 601 GGCCATCTTTTCCCTCATGGTATTATGTCCTTGAAGAGCGCTTTGAGAGATCTAGTGGGC 660
Oy 661 TTTCTTTCTGGGTTTGGGCAATTCAGTTCATGTCATGTCATATTCTATCATTAATTGTA 720
Db 661 TTTCTTTCTGGGTTTGGGCAATTCAGTTCATGTCATGTCATATTCTATCATTAATTGTA 720
Oy 721 TAACGGTTTCAAAACAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGAAAT 780
Db 721 TAACGGTTTCAAAACAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGAAAT 780
Oy 781 AGCCACGGGAGATCTCCAGACCAATCTCCATGTTTCCACAGCTCTCCAGCCAAACC 840
Db 781 AGCCACGGGAGATCTCCAGACCAATCTCCATGTTTCCACAGCTCTCCAGCCAAACC 840
Oy 841 AAATAGCCCTGCTATAGTATGATCATCTCGGCGCTTGAAGCTTGTCCCTCTTAATG 900
Db 841 AAATAGCCCTGCTATAGTATGATCATCTCGGCGCTTGAAGCTTGTCCCTCTTAATG 900
Oy 901 TTCTTTAATCAGATTAAGTGCCTGGAAGCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960
Db 901 TTCTTTAATCAGATTAAGTGCCTGGAAGCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960
Oy 961 TTGCTAAGTGAATTAATGAGTGTGTTTTTCCGTATTAAGCAAAATTAATTTAAAAATG 1020
Db 961 TTGCTAAGTGAATTAATGAGTGTGTTTTTCCGTATTAAGCAAAATTAATTTAAAAATG 1020
Oy 1021 AAAAGTT 1027
Db 1021 AAAAGTT 1027

```

```

RESULT 2
LOCUS AF151103 1027 bp mRNA linear PRI 03-NOV-2000
DEFINITION Homo sapiens TCRgamma alternate reading frame protein (TCRG) mRNA,
ACCESSION AF151103
VERSION AF151103.1 GI:5758136
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

```

REFERENCE
1 (sites)
AUTHORS Essand,M., Vasmatazis,G., Brinkmann,U., Dury,P., Lee,B. and Pastan,I.
TITLE High expression of a specific T-cell receptor gamma transcript in epithelial cells of the prostate
PROC. Natl. Acad. Sci. U.S.A. 96 (16), 9287-9292 (1999)
JOURNAL
PUBMED 10430935
REFERENCE
2 (bases 1 to 1027)
AUTHORS Wolfgang,C.D., Essand,M., Vincent,J.J., Lee,B. and Pastan,I.
TITLE TRP: a nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor gamma chain locus
PROC. Natl. Acad. Sci. U.S.A. 97 (17), 9437-9442 (2000)
JOURNAL
PUBMED 10931945
REFERENCE
3 (bases 1 to 1027)
AUTHORS Essand,M., Vasmatazis,G., Brinkmann,U., Dury,P., Lee,B. and Pastan,I.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-1999) Division of Basic Sciences, Laboratory of Molecular Biology, National Institutes of Health, National Cancer Institute, 37 Convent Drive, Bethesda, MD 20892, USA
FEATURES
source
1..1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/note="transcript from an unarranged TCRg locus"
1..1027
/gene="TCRG"
1..60
/gene="TCRG"
/notes="gamma 1.2"
61..390
/gene="TCRG"
/notes="gamma 1; corresponds to exon I"
74..250
/gene="TCRG"
/notes="TRAP"
/codon_start=1
/product="TCRgamma alternate reading frame protein"
/protein_id="AA929337.1"
/db_xref="GI:11093514"
/translation="MQMPPSPLEFFQLKQSSRLRLEHFFVLFNFSMLRLRYGKK RRAIRFMDPRGRP"
391..438
/gene="TCRG"
/notes="gamma 1; corresponds to exon II"
439..579
/gene="TCRG"
/notes="gamma 1; corresponds to exon III"

```

ORIGIN

```

Query Match 100.0%; Score 1027; DB 8; Length 1027;
Best Local Similarity 100.0%; Pred. No. 3; Se-253; Indels 0; Gaps 0;
Matches 1027; Conservative 0; Mismatches 0;

```

```

Oy 1 GGGCAAGAGTTGGGCAAAAATCAAGTATTTGTCCTGGACAAAGCTTATCATTA 60
Db 1 GGGCAAGAGTTGGGCAAAAATCAAGTATTTGTCCTGGACAAAGCTTATCATTA 60
Oy 61 GATAAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCT 120
Db 61 GATAAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCT 120
Oy 121 GAAACAAAGCTCCGAAAGCTGGAACATCTTTGTCTTCTTGGAATTTTCCCTGAT 180
Db 121 GAAACAAAGCTCCGAAAGCTGGAACATCTTTGTCTTCTTGGAATTTTCCCTGAT 180
Oy 181 GTTATTAAAGATACATTGGCAAGAAAAGACACGATTTCTGGGATCCCGAGAGGG 240
Db 181 GTTATTAAAGATACATTGGCAAGAAAAGACACGATTTCTGGGATCCCGAGAGGG 240

```

QY 241 AACACATGAAGACTAAGACACATACATGAATTTAGTGTAAACGATGCGAGAAAG 300
 DB 241 AACACATGAAGACTAAGACACATACATGAATTTAGTGTAAACGATGCGAGAAAG 300
 QY 301 TCACCTGACAAAGACACAGATGATGTCAGACATGAGATATATAAAAAAGAGTTGAT 360
 DB 301 TCACCTGACAAAGACACAGATGATGTCAGACATGAGATATATAAAAAAGAGTTGAT 360
 QY 361 CAAGAAATTTATCTTCTCCATTAAGACGGATGTCATCAATGATATCCCAAGACAT 420
 DB 361 CAAGAAATTTATCTTCTCCATTAAGACGGATGTCATCAATGATATCCCAAGACAT 420
 QY 421 TGTTCAAAAGATGCAATGATATACACTACTGCTGAGCTCAAAACCTCTGCATATTAC 480
 DB 421 TGTTCAAAAGATGCAATGATATACACTACTGCTGAGCTCAAAACCTCTGCATATTAC 480
 QY 481 ATGTACTCTCTCTGCTCTCAAGAGTGTGATCTATTTTGCATCATACCTGCTGTG 540
 DB 481 ATGTACTCTCTCTGCTCTCAAGAGTGTGATCTATTTTGCATCATACCTGCTGTG 540
 QY 541 CTTAGAAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600
 DB 541 CTTAGAAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600
 QY 601 GGCCATCTTTTCTCATCGGTTATGTCCTTAGAAGGCTCTTCAAGGATCTAGTTGGGC 660
 DB 601 GGCCATCTTTTCTCATCGGTTATGTCCTTAGAAGGCTCTTCAAGGATCTAGTTGGGC 660
 QY 661 TTTCTTCTGAGTTTGGGCAATTCAGTTCTCATGTCATGTCATTTCTATCATTTATTTGA 720
 DB 661 TTTCTTCTGAGTTTGGGCAATTCAGTTCTCATGTCATGTCATTTCTATCATTTATTTGA 720
 QY 721 TAAACGTTTTCAAACCAAGTGGGCAACAGAGAACCTCACTCTGTAATTAAGAGGAAT 780
 DB 721 TAAACGTTTTCAAACCAAGTGGGCAACAGAGAACCTCACTCTGTAATTAAGAGGAAT 780
 QY 781 AGCCAGCGGATCTCCAGACCAATCTCTCATGTTTCCAGAGCTCTCCAGCAAGCC 840
 DB 781 AGCCAGCGGATCTCCAGACCAATCTCTCATGTTTCCAGAGCTCTCCAGCAAGCC 840
 QY 841 AAATAGCGCTGCTATAGTATAGACATCTGCGGCTTCTAGCCTTGTCCCTCTTTAGTG 900
 DB 841 AAATAGCGCTGCTATAGTATAGACATCTGCGGCTTCTAGCCTTGTCCCTCTTTAGTG 900
 QY 901 TTTCTTAATCAGATTAATGCTGCTGAGACCTTTTCAATTTTAAACGCGCTGAAGAGTTCT 960
 DB 901 TTTCTTAATCAGATTAATGCTGCTGAGACCTTTTCAATTTTAAACGCGCTGAAGAGTTCT 960
 QY 961 TTGCTAGTTGAATTAATGCTGCTTTTCCGTAATTAAGCAAAATTAATTAATAAATG 1020
 DB 961 TTGCTAGTTGAATTAATGCTGCTTTTCCGTAATTAAGCAAAATTAATTAATAAATG 1020
 QY 1021 AAAAGTT 1027
 DB 1021 AAAAGTT 1027

RESULT 3
 LOCUS AX074415 1027 bp DNA linear PAT 06-FEB-2001
 DEFINITION Sequence 13 from Patent WO0104309.
 ACCESSION AX074415
 VERSION AX074415.1 GI:12710553
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1
 AUTHORS Patsan, I., Bessand, M., Lee, B., Vaarmatzis, G. and Wolfgang, C.
 TITLE T-cell receptor γ (g) alternate reading frame protein, (tarp) and
 uses thereof

JOURNAL Patent: WO 0104309-A 13 18-JAN-2001;
 UNITED STATES GOVERNMENT (US)
 FEATURES Location/Qualifiers
 source 1..1027
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 74..250
 CDS
 /note="unnamed protein product; Coding region for PS-TCR
 gamma 1 polypeptide (TARF)"
 /codon_start=1
 /protein_id="CAC28473.1"
 /db_xref="GI:12710554"
 /translation="MQMFPSPPLFFLLQKSSRLRHTFVFLNFSMLRXYGKK
 RRAIRFMDRRGTP"
 247..582
 CDS
 /note="unnamed protein product; Coding region for PS-TCR
 gamma 2 polypeptide (deduced amino acid sequence not
 displayed along with DNA sequence, due to overlapping
 CDS's)"
 /codon_start=1
 /protein_id="CAC28474.1"
 /db_xref="GI:12710555"
 /translation="MKTNDTYMKFSMLTVPEKSLDKRHCIVRHNNKNGVDQELIFP
 PKITDITDPDKNCSKDANDTLILQLNTISAYMYLILLLKSVYFAIITCCLLRRT
 AFCNGEKS"
 ORIGIN
 Query Match 99.8%; Score 1025.4; DB 6; Length 1027;
 Best Local Similarity 99.9%; Pred. No. 9.1e-253;
 Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCAAGAGTTGGGCAAAATTAAGATTTTGTGTCCTGGGCAACAGCTTATCATTAACA 60
 DB 1 GGGCAAGAGTTGGGCAAAATTAAGATTTTGTGTCCTGGGCAACAGCTTATCATTAACA 60
 QY 61 GATTAACAATCTGATGATGATGTTTCCCAAGCCCACTATTCTTCTCTCAATTGCT 120
 DB 61 GATTAACAATCTGATGATGATGTTTCCCAAGCCCACTATTCTTCTCTCAATTGCT 120
 QY 121 GAAACAAGCTCCAGAGGCTGGACATACCTTGTCTTCTGAGAATTTTCCCTGAT 180
 DB 121 GAAACAAGCTCCAGAGGCTGGACATACCTTGTCTTCTGAGAATTTTCCCTGAT 180
 QY 181 GTTATTAAGATACATTGGCAAGAAAGAGCAACAGATTTCTGGATCCCGAGGGG 240
 DB 181 GTTATTAAGATACATTGGCAAGAAAGAGCAACAGATTTCTGGATCCCGAGGGG 240
 QY 241 AACACATGAAGCTTAAGACACATACATGAATTTAGCTGTTAACGTCGAGAAAG 300
 DB 241 AACACATGAAGCTTAAGACACATACATGAATTTAGCTGTTAACGTCGAGAAAG 300
 QY 301 TCACCTGACAAAGACACAGATGATGTCAGACATGAGATATATAAAAAAGAGTTGAT 360
 DB 301 TCACCTGACAAAGACACAGATGATGTCAGACATGAGATATATAAAAAAGAGTTGAT 360
 QY 361 CAAGAAATTTATCTTCTCCATTAAGACGGATGTCATCAATGATATCCCAAGACAT 420
 DB 361 CAAGAAATTTATCTTCTCCATTAAGACGGATGTCATCAATGATATCCCAAGACAT 420
 QY 421 TGTTCAAAAGATGCAATGATATACACTACTGCTGAGCTCAAAACCTCTGCATATTAC 480
 DB 421 TGTTCAAAAGATGCAATGATATACACTACTGCTGAGCTCAAAACCTCTGCATATTAC 480
 QY 481 ATGTACTCTCTCTGCTCTCAAGAGTGTGATCTATTTTGCATCATACCTGCTGTG 540
 DB 481 ATGTACTCTCTCTGCTCTCAAGAGTGTGATCTATTTTGCATCATACCTGCTGTG 540
 QY 541 CTTAGAAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600
 DB 541 CTTAGAAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600
 QY 601 GGCCATCTTTTCTCATCGGTTATGTCCTTAGAAGGCTCTTCAAGGATCTAGTTGGGC 660

```

Db      601 GGCACATCTTTTCATCGGTATGTCCTTGAAGAGCCTCTTGAGAGATCATGTTGGGC 660
Oy      661 TTTCTTCTGGGTTTGGGCACTTTGAGTTTCATGTGTACTATTTATCATTTATGTA 720
Db      661 TTTCTTCTGGGTTTGGGCACTTTGAGTTTCATGTGTACTATTTATCATTTATGTA 720
Oy      721 TAAAGGTTTTCACCAAGTGGGACACAGAGAACTCTCTGTATTAATACATGAGAAAT 780
Db      721 TAAAGGTTTTCACCAAGTGGGACACAGAGAACTCTCTGTATTAATACATGAGAAAT 780
Oy      781 AGCCAGCGGAGATCTCCAGACCAATCTCCATGTTTCCACAGTCTCTCCAGCAACCC 840
Db      781 AGCCAGCGGAGATCTCCAGACCAATCTCTCCATGTTTCCACAGTCTCTCCAGCAACCC 840
Oy      841 AAATAGCCCTGCTCTAGTGTAGATCCTGGGCTTTCAGCCTTGTCCCTCTTAAATG 900
Db      841 AAATAGCCCTGCTCTAGTGTAGATCCTGGGCTTTCAGCCTTGTCCCTCTTAAATG 900
Oy      901 TTTCTTAAATCAGATTAAGTCCCTGGAAGCCTTTCATTTTACAAGCCCTGAGAGCTTCT 960
Db      901 TTTCTTAAATCAGATTAAGTCCCTGGAAGCCTTTCATTTTACAAGCCCTGAGAGCTTCT 960
Oy      961 TTGCTAGTTGAATATGTGTGTGTGTTTTCCTGTAATAGCAAAATTTAAAAAATG 1020
Db      961 TTGCTAGTTGAATATGTGTGTGTGTTTTCCTGTAATAGCAAAATTTAAAAAATG 1020
Oy      1021 AAAAGTT 1027
Db      1021 AAAAGTT 1027

```

```

RESULT 4
LOCUS      BC072396          1825 bp      mRNA      linear      PRI 30-JUN-2004
DEFINITION Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone
ACCESSION   BC072396
VERSION     BC072396.1 GI:47939858
KEYWORDS
SOURCE      MGC.
ORGANISM    Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE   1 (bases 1 to 1825)
AUTHORS    Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Scheet C.F., Bhat N.K.,
Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Helel F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Scheet T.E., Brownstein M.J., Umed T.B., Toshiki S.,
Carroll J.P., Prange C., Kana S.S., Loquellano N.A., Peters G.J.,
Abrahamson R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulvik S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S.,
Sanchez A., Whitting M., Madan A., Young A.C., Scherchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myer R.M.,
Butcherfield Y.S., Krzywicki M.I., Skalka U., Smalhus D.E.,
Scherer A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
PUBMED     12477932
AUTHORS    Strausberg R.
TITLE      Direct Submission
JOURNAL    Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

```

REMARK

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland.
 Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nrl.nih.gov
 Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Masketti, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,
 Teurisson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 1/4 Row: c Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

SOURCE

1..1825
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:90486 IMAGE:5755352"
 /issue_type="Blood, adult leukocytes"
 /clone_id="NIH_MGC_118"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..1825

gene

CDS

/gene="TRGV9"
 /note="synonyms: TCRGV9, V2"
 /db_xref="GeneID:6983"
 /db_xref="IMGT/IRGM:TRGV9"
 378..1325
 /gene="TRGV9"
 /codon_start=1
 /product="TRGV9 protein"
 /protein_id="AAH72396.1"
 /db_xref="GI:47939859"
 /db_xref="GeneID:6983"
 /translation="MLSLHSTLAVALGCVGAGHLEDPQISTKTLSTKLBCV
 VSGTIGATSVVYWRERPGVIGVIGLVSIGVQYRKSGPSGFRFDRIPERTSL
 TTHNVEKQDIATYICALMEVFGELGKIKIRFGPKLITDKDLADVSPKPTFLS
 IAEKVLQKATYICLLEKFPDVKIKHMERKSTLIGSDEGNTKNDYMKSWLT
 VPEKSLDKERCIYHNNKNGVDQEIIFPIKTDVITMPKDCSDANDTLLOLT
 NTSAYVYVYLLLLKSVYVPAITCCLLRATFCNGSKS"

ORIGIN

Query Match 98.3%; Score 1010; DB 8; Length 1825;
 Best Local Similarity 99.5%; Pred. No. 8e-249;
 Matches 1013; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Oy 7 GAGTTGGGCAAAAATCAAGCTTTTGGTCCCGAACAAGCTTATCATGAGATTA 66
 Db 750 GAGTTGGGCAAAAATCAAGCTTTTGGTCCCGAACAAGCTTATCATGAGATTA 809
 Oy 67 CAACTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGAACA 126
 Db 810 CAACTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGAACA 869
 Oy 127 AAGCTCAGAAAGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCGTGATTA 186

D	b		870	AAGCTCAGAAGGCTGGAAACATACCTTTGTCTTCTTGAGAAATTTCCTCGATGTTATT	929
O	y		187	AAGATACATTGGCAAGAAAAGAAAGACAACGCATTTCTGGGATTCGAGAGGGGAACC	246
D	b		930	AAGATACATTGGGAAAGAAAAGAAAGAACACGATTCCTGGGATCCGAGAGGGAAACC	989
O	y		247	ATGAAGAATAAGCAACATACATGAAATTAATAGCTGGTTAAGGTCGCCAGAAAATGCACTG	306
D	b		990	ATGAAGACTTAATGACACATACATGAAATTTAGCTGGTTAAGCGTGCCAGAAAATGCACTG	1049
O	y		307	GACAAAGAACACAGATGTATGTCGACAGCATGAGATATATAAAAAACGAGTTGATCAA	366
D	b		1050	GACAAAGAACACAGATGTATGTCGACAGCATGAGATATATAAAAAACGAGTTGATCAA	1109
O	y		367	ATTATCTTTCTCCATTAAGAAGGATGTCATCACATGATGCCCAAGACAAATTTGTCA	426
D	b		1110	ATTATCTTTCTCCATTAAGAAGGATGTCATCACATGATGCCCAAGACAAATTTGTCA	1169
O	y		427	AAAGATGCAATGATGATCACTGCTGAGTCCAACACCTCGCATATTACATGATC	486
D	b		1170	AAAGATGCAATGATGATCACTGCTGAGTCCAACACCTCGCATATTACATGATC	1229
O	y		487	CTCCTCCTGCTCTCAAGAGTGGTCTAATTTGCCATCATCACTGCTGCTCTTAGA	546
D	b		1230	CTCCTCCTGCTCTCAAGAGTGGTCTAATTTGCCATCATCACTGCTGCTCTTAGA	1289
O	y		547	AGAAAGGCTTTCTGCTGATGAGAGAAATCATPAACAGAGGTTGGCAAGAGGCCAT	606
D	b		1290	AGAAAGGCTTTCTGCTGATGAGAGAAATCATPAACAGAGGTTGGCAAGAGGCCAT	1349
O	y		607	CTTTTCTCATACGGTTAATGTCCCTAGAAAGGCTCTTGAAGATCTAATGGGCTTTCTT	666
D	b		1350	CTTTTCTCATACGGTTAATGTCCCTAGAAAGGCTCTTGAAGATCTAATGGGCTTTCTT	1409
O	y		667	TCTGGGTTTGGGCACTTTCAGTGTCTCATGTGNGTACTATTCATCATATTGTTAAACGG	726
D	b		1410	TCTGGGTTTGGGCACTTTCAGTGTCTCATGTGNGTACTATTCATCATATTGTTAAACGG	1469
O	y		727	TTTTCAAAACAGTGGGCAACAGAGAACTCACTCTGTATTAACAGATGAGAAATAGCCAC	786
D	b		1470	TTTTCAAAACAGTGGGCAACAGAGAACTCACTCTGTATTAACAGATGAGAAATAGCCAC	1529
O	y		787	GGCGATCTCCAGCACCAATCTCTCATGTGTTTCCACAGCTCTTCAAGCCAAATAG	846
D	b		1530	GGCGATCTCCAGCACCAATCTCTCATGTGTTTCCACAGCTCTTCAAGCCAAATAG	1589
O	y		847	CGCGTCGATATGTGTAGACATCCGCGGCTCTAGCCCTGTGCCCCCTTATGTGTTCTT	906
D	b		1590	CGCGTCGATATGTGTAGACATCCGCGGCTCTAGCCCTGTGCCCCCTTATGTGTTCTT	1649
O	y		907	AATCAGATTAACGTGCTCGAAGACCTTTCAATTTTACACGCCCTGAAGACATCTTCTTGCTA	966
D	b		1650	AATCAGATTAACGTGCTCGAAGACCTTTCAATTTTACACGCCCTGAAGACATCTTCTTGCTA	1709
O	y		967	GTTGAAATTAATGTGTGTGTTTTCCTGTATTAAGCAAAATTAATTTAAAAAATGAAA	1024
D	b		1710	GTTGAAATTAATGTGTGTGTTTTCCTGTATTAAGCAAAATTAATTTAAAAAATGAAA	1767
RESULT 5					
		BC062761	1054 bp	mRNA	linear PRI 16-DEC-20030
		LOCUS			
		DEFINITION	Homo sapiens T cell receptor gamma variable 9, mRNA (CDNA clone		
		VERSION	IM654:4248993), partial cds.		
		KEYWORDS	BC062761.1 GI:38540986		
		SOURCE			
		ORGANISM	Homo sapiens (human)		
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
			Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
			Homnidae; Homo.		
		REFERENCE	1 (bases 1 to 1054)		

REMARK	COMMENT	JOURNAL	TITLE	AUTHORS
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk Email: cgabds-remail.nih.gov Tissue Procurement: CLONTECH CNA Library Preparation: CLONTECH Laboratories, Inc. CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@cgsc.bc.ca	Submitted (25-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932 2 (bases 1 to 1054)	Krausenberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Stansberg, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavani, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunnarsson, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketterman, M., Madan, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk Email: cgabds-remail.nih.gov Tissue Procurement: CLONTECH CNA Library Preparation: CLONTECH Laboratories, Inc. CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@cgsc.bc.ca	Submitted (25-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932 2 (bases 1 to 1054)	Krausenberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Stansberg, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavani, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunnarsson, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketterman, M., Madan, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk Email: cgabds-remail.nih.gov Tissue Procurement: CLONTECH CNA Library Preparation: CLONTECH Laboratories, Inc. CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@cgsc.bc.ca	Submitted (25-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932 2 (bases 1 to 1054)	Krausenberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Stansberg, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavani, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunnarsson, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketterman, M., Madan, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk Email: cgabds-remail.nih.gov Tissue Procurement: CLONTECH CNA Library Preparation: CLONTECH Laboratories, Inc. CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@cgsc.bc.ca	Submitted (25-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932 2 (bases 1 to 1054)	Krausenberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Stansberg, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavani, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunnarsson, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketterman, M., Madan, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk Email: cgabds-remail.nih.gov Tissue Procurement: CLONTECH CNA Library Preparation: CLONTECH Laboratories, Inc. CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@cgsc.bc.ca	Submitted (25-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932 2 (bases 1 to 1054)	Krausenberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Stansberg, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavani, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunnarsson, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketterman, M., Madan, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
NIH-MGC Project URL:				

```
misc_feature
      74..364
      /gene="TRGV9"
      /note="IG; Region: Immunoglobulin domain constant region
      subfamily"
      /db_xref="CDD:c000098"

ORIGIN
Query Match      98.2%; Score 1008.8; DB 8; Length 1054;
Best Local Similarity 99.8%; Pred. No. 1.7e-248;
Matches 1010; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

      16 AAAAAAAAAAGGATTTTGTCCCGGAAAGCTTATCATTAAGATTAACACTTGAT 75
      11 AAAAAAAAAAGGATTTTGTCCCGGAAAGCTTATCATTAAGATTAACACTTGAT 70
      76 GCAAGTGTTCCTCCCAAGCCCATATTTTCTCTCTCAATTCGTAAGAAAGCTCCAG 135
      71 GCAAGTGTTCCTCCCAAGCCCATATTTTCTCTCTCAATTCGTAAGAAAGCTCCAG 130
      136 AAGGCTGGAACATACCTTTGTCTCTTGAAGAAATTTTCCCTGATGTTATTAGATACAT 195
      131 AAGGCTGGAACATACCTTTGTCTCTTGAAGAAATTTTCCCTGATGTTATTAGATACAT 190
      196 TGGCAAGAAAAGAGAGCAACGATTTGGGATCCGAGAGGGGAAACCATGAGACT 255
      191 TGGCAAGAAAAGAGAGCAACGATTTGGGATCCGAGAGGGGAAACCATGAGACT 250
      256 AAGCAACATCATGAATTTAGCTGTTAAGCGGCAAGAAAGTCACTGGAACAAAGAA 315
      251 AAGCAACATCATGAATTTAGCTGTTAAGCGGCAAGAAAGTCACTGGAACAAAGAA 310
      316 CACAGATGATCTGCAGACATGAGATATATAAAGAGATGATCAAGAAATTAATCTTT 375
      311 CACAGATGATCTGCAGACATGAGATATATAAAGAGATGATCAAGAAATTAATCTTT 370
      376 CCTCCATTAAGACGAGATGTCATCAATGATGCCAAGACAAATGTTCAAAAGATCA 435
      371 CCTCCATTAAGACGAGATGTCATCAATGATGCCAAGACAAATGTTCAAAAGATCA 430
      436 AATGATACATGCTGCTGCAAGCTCAAAACCTCTGATATTAATGATGATCTCTCTG 495
      431 AATGATACATGCTGCTGCAAGCTCAAAACCTCTGATATTAATGATGATCTCTCTG 490
      496 CTCTCAAGAGTGTGTCTATTTTGCATCATCATCTGCTGTCTGCTTGAAGAACGCT 555
      491 CTCTCAAGAGTGTGTCTATTTTGCATCATCATCTGCTGTCTGCTTGAAGAACGCT 550
      556 TTCTGCTCAATGAGAGAAATCATTAACAGAGGTGGCAAGAGAGGCAATCTTTCTC 615
      551 TTCTGCTCAATGAGAGAAATCATTAACAGAGGTGGCAAGAGAGGCAATCTTTCTC 610
      616 ATCGGTATTTGCTCTAGAGGCTCTTGAAGATCTAGTGGGCTTTCTTTCTGGGTTT 675
      611 ATCGGTATTTGCTCTAGAGGCTCTTGAAGATCTAGTGGGCTTTCTTTCTGGGTTT 670
      676 GGGCCATTTCAATTCTCATGTTGTACTATTCTATCATTAATGTTAAACGTTTCAAC 735
      671 GGGCCATTTCAATTCTCATGTTGTACTATTCTATCATTAATGTTAAACGTTTCAAC 730
      736 CAGTGGGCAACAGAGAACTCATCTGTATTAACAATGAGGAATAGCCAGGATCTC 795
      731 CAGTGGGCAACAGAGAACTCATCTGTATTAACAATGAGGAATAGCCAGGATCTC 790
      796 CAGCAACAATCTCTCAAGTTTTCACAGCTCTCTCAAGCAACCAAAATAGGCTCTGTA 855
      791 CAGCAACAATCTCTCAAGTTTTCACAGCTCTCTCAAGCAACCAAAATAGGCTCTGTA 850
```

```
QY      856 TAGTGAGACATCTGCGGCTTCTAGCCTTGTCCCTCTTATGTTCTTTAATCAATA 915
DB      851 TAGTGAGACATCTGCGGCTTCTAGCCTTGTCCCTCTTATGTTCTTTAATCAATA 910
QY      916 ACTGCTCGAAGCCTTTCATTTTACAGCCCTGAAGAGCTTCTTGTCTAGTTGAATTA 975
DB      911 ACTGCTCGAAGCCTTTCATTTTACAGCCCTGAAGAGCTTCTTGTCTAGTTGAATTA 970
QY      976 TGTGCTGTTTTCCTTAATTAAGCAAAATTAATTTAAAAATGAAGAGTT 1027
DB      971 TGTGCTGTTTTCCTTAATTAAGCAAAATTAATTTAAAAATGAAGAGTT 1022

RESULT 6
CS032596      1530 bp      DNA      Linear      PAT 10-MAR-2005
LOCUS
DEFINITION
Sequence 2102 from Patent WO2005016962.
ACCESSION
CS032596
VERSION
CS032596.1 GI:60732060
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1
Abbas, A., Clark, H., Ouyang, W., Williams, M. P., Wood, W. I. and Wu, T. D.
Compositions and methods for the treatment of immune related
diseases
Patent: WO 2005016962-A 2102 24-FEB-2005;
Genentech, Inc. (US)
JOURNAL
Location/Qualifiers
FEATURES
source
1..1530
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      95.8%; Score 984; DB 6; Length 1530;
Best Local Similarity 98.0%; Pred. No. 4e-242;
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

      12 GGGCAAAAAATCAAGTATTTGTGTCCTCCGAAACAAAGCTTATCATTAAGATTAACA 71
      515 GGGAAATTAATTAAGAACTCTTGGCAGTGGAAACAACCTGTGTCAAGATTAACAAC 574
      72 TGATGCAAGATTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGGAAGAGCT 131
      575 TGATGCAAGATTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGGAAGAGCT 634
      132 CCAGAGGCTGGAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGAT 191
      635 CCAGAGGCTGGAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGAT 694
      192 ACATTGGCAAGAAAAGAGCAACAGATTTCTGGATCCAGAGAGGGAAACCATGAA 251
      695 ACATTGGCAAGAAAAGAGCAACAGATTTCTGGATCCAGAGAGGGAAACCATGAA 754
      252 GACTTAAGCAACATACATGAATTAATTAAGTGTAAACGTTGCAAGAAAGTCACTGGA 311
      755 GACTTAAGCAACATACATGAATTAATTAAGTGTAAACGTTGCAAGAAAGTCACTGGA 814
      312 AGAACAAGATGTATCTGCAGACATGAGATTAATTAAGAGATGATCAAGAAATTAAT 371
      815 AGAACAAGATGTATCTGCAGACATGAGATTAATTAAGAGATGATCAAGAAATTAAT 874
      372 CTTCCTCCATTAAGAGAGATGTCATCAATGATCCCAAGACAAATTTTCAAAAAGA 431
      875 CTTCCTCCATTAAGAGAGATGTCATCAATGATCCCAAGACAAATTTTCAAAAAGA 934
      432 TGCAATGATACATGCTGCTGCAAGTCAAAACCTCTGCAATTAATGATGATCTCT 491
```



```

Db      935 TGCAAATGATACACTACTGCTGACAGCTCACAAACACTCTGCAATTATACATGATACCTCT 994
Qy      492 CCGCTCCCTCAAGAGTGTGTCTATTTTGGCATTATACCTGCTGTCTGCTTAGAAGAAC 551
Db      995 CCGTCTCCTCAAGAGTGTGTCTATTTTGCATCATACCTGCTGTCTGCTTAGAAGAAC 1054
Qy      552 GGGCTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGACACAGAGAGCCATCTTTT 611
Db      1055 GGGCTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGACACAGAGAGCCATCTTTT 1114
Qy      612 CTTCAATGGTATATGCTCTTGAAGAGCGTCTTGTAGAGATCTAGTGGGCTTTCTTTCTGG 671
Db      1115 CTTCAATGGTATATGCTCTTGAAGAGCGTCTTGTAGAGATCTAGTGGGCTTTCTTTCTGG 1174
Qy      672 GTTTGGGCGCATTTTCAGTTCTCATGTGTGTACTATTTATCATTTATTTATTAACGGTTTTC 731
Db      1175 GTTTGGGCGCATTTTCAGTTCTCATGTGTGTACTATTTATCATTTATTTATTAACGGTTTTC 1234
Qy      732 AAACCAAGTGGGACACAGAGAACCTCACTCTGTATATACATGAGAAATAGCCACGGCGA 791
Db      1235 AAACCAAGTGGGACACAGAGAACCTCACTCTGTATATACATGAGAAATAGCCACGGCGA 1294
Qy      792 TCTCCAGACCAATCTCTCCATGTTTTTCAACAGCTCTCCAGCCCAACCAATAGCGCCT 851
Db      1295 TCTCCAGACCAATCTCTCCATGTTTTTCAACAGCTCTCCAGCCCAACCAATAGCGCCT 1354
Qy      852 GCTATAGTGAACATCTCTGGGCTTCTAGCCTTGTCTCTCTTATGTTCTTTATCA 911
Db      1355 GCTATAGTGAACATCTCTGGGCTTCTAGCCTTGTCTCTCTTATGTTCTTTATCA 1414
Qy      912 GATPACGCGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGCTCTTGTGAGTTGA 971
Db      1415 GATPACGCGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGCTCTTGTGAGTTGA 1474
Qy      972 ATTATGTGTGTGTTTTTCCGTATATAGCAAAATTAATTTAAAAATGAAAAGTT 1027
Db      1475 ATTATGTGTGTGTTTTTCCGTATATAGCAAAATTAATTTAAAAATGAAAAGTT 1530

RESULT 7
CS041548      1530 bp      DNA      linear      PAT 22-MAR-2005
LOCUS      CS041548
DEFINITION      Sequence 2102 from Patent WO2005019258.
ACCESSION      CS041548
VERSION      CS041548.1 GI:61848985
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Abbae,A., Clark,H., Ouyang,W., Williams,P.M., Wood,W.I. and Wu,T.D.
AUTHORS      Compositions and methods for the treatment of immune related
TITLE      diseases
JOURNAL      Patent: WO 2005019258-A 2102 03-MAR-2005;
GENE/PROTEIN      Genentech, Inc. (US)
FEATURES
SOURCE
1. 1530
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      95.8%; Score 984; DB 6; Length 1530;
Best Local Similarity 98.0%; Pred. No. 4e-242;
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy      12 GGGCAAAAATCAAGTATTGTCCTCCGGAACAAGCTTATCATTAACAGATTAACACT 71
Db      515 GGGAAATTAATGAAGAACTCTTGGACAGTGGAAACAACATCTGTTCTCAGATTAACAAC 574
Qy      72 TGATGAGATGTTTCCCAAGCCACATATTTTCTTCTTCAATGCTGAAAACAAGCT 131

```

```

Db      575 TGATGAGATGTTTCCCAAGCCACATATTTTCTTCTTCAATGCTGAAAACAAGCT 634
Qy      132 CCAAGAGCTGGAACATACCTTTGTTCTTGTGAAAATTTTCCGATGTTATTAAGAT 191
Db      635 CCAAGAGCTGGAACATACCTTTGTTCTTGTGAAAATTTTCCGATGTTATTAAGAT 694
Qy      192 ACAATGGCAAGAAAAGAGCAACAAGATTCGGGATCCAGAGGGGAAACCATGAA 251
Db      695 ACAATGGCAAGAAAAGAGCAACAAGATTCGGGATCCAGAGGGGAAACCATGAA 754
Qy      252 GACTTAACACATACATGAAATTTAGCTGTGTTAACGGTCCAGAAAAGTCACTGCA 311
Db      755 GACTTAACACATACATGAAATTTAGCTGTGTTAACGGTCCAGAAAAGTCACTGCA 814
Qy      312 AGAACACAGATGATGTGTGACATAGAAATTAATAAAAGGAGTTGATTAAGAAATTA 371
Db      815 AGAACACAGATGATGTGTGACATAGAAATTAATAAAAGGAGTTGATTAAGAAATTA 874
Qy      372 CTTTCTCCAAATTAAGACGAGATGTCATCAATGAGATCCCAAGACATTTGTTCAAGA 431
Db      875 CTTTCTCCAAATTAAGACGAGATGTCATCAATGAGATCCCAAGACATTTGTTCAAGA 934
Qy      432 TGCAAATGATACACTACTGCTGACGCTCAACAAACACTCTGCATATTATACATGACCTCT 491
Db      935 TGCAAATGATACACTACTGCTGACGCTCAACAAACACTCTGCATATTATACATGACCTCT 994
Qy      492 CTTGCTCTCAAGAGTGTGTCTATTTTGGCATCATCACTGCTGTCTGTGTTAAGAAC 551
Db      995 CTTGCTCTCAAGAGTGTGTCTATTTTGGCATCATCACTGCTGTCTGTGTTAAGAAC 1054
Qy      552 GGGCTTCTGCTGAAATGAGAGAAATCATPAACAGAGGTGGCAACAAGAGGCCATCTTTT 611
Db      1055 GGGCTTCTGCTGAAATGAGAGAAATCATPAACAGAGGTGGCAACAAGAGGCCATCTTTT 1114
Qy      612 CTTCAATGGTATATGCTCTTGAAGAGCGTCTTGTAGAGATCTAGTGGGCTTTCTTCTGG 671
Db      1115 CTTCAATGGTATATGCTCTTGAAGAGCGTCTTGTAGAGATCTAGTGGGCTTTCTTCTGG 1174
Qy      672 GTTTGGGCGCATTTTCAGTTCTCATGTGTGTACTATTTATCATTTATTTATTAACGGTTTTC 731
Db      1175 GTTTGGGCGCATTTTCAGTTCTCATGTGTGTACTATTTATCATTTATTTATTAACGGTTTTC 1234
Qy      732 AAACCAAGTGGGACACAGAGAACCTCACTCTGTATATACATGAGAAATAGCCACGGCGA 791
Db      1235 AAACCAAGTGGGACACAGAGAACCTCACTCTGTATATACATGAGAAATAGCCACGGCGA 1294
Qy      792 TCTCCAGACCAATCTCTCCATGTTTTTCAACAGCTCTCCAGCCCAACCAATAGCGCCT 851
Db      1295 TCTCCAGACCAATCTCTCCATGTTTTTCAACAGCTCTCCAGCCCAACCAATAGCGCCT 1354
Qy      852 GCTATAGTGAACATCTCTGGGCTTCTAGCCTTGTCTCTCTTATGTTCTTTATCA 911
Db      1355 GCTATAGTGAACATCTCTGGGCTTCTAGCCTTGTCTCTCTTATGTTCTTTATCA 1414
Qy      912 GATPACGCGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGCTCTTGTGAGTTGA 971
Db      1415 GATPACGCGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGCTCTTGTGAGTTGA 1474
Qy      972 ATTATGTGTGTGTTTTTCCGTATATAGCAAAATTAATTTAAAAATGAAAAGTT 1027
Db      1475 ATTATGTGTGTGTTTTTCCGTATATAGCAAAATTAATTTAAAAATGAAAAGTT 1530

RESULT 8
HUMTCRGA      1530 bp      mRNA      linear      PRI 14-JAN-1995
LOCUS      HUMTCRGA
DEFINITION      Human T-cell receptor gamma chain VCI1-CII-CIII region mRNA,
complete cds.
ACCESSION      M16768
VERSION      M16768.1 GI:339399
KEYWORDS      T-cell receptor.
SOURCE      Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1 (bases 1 to 1530)
AUTHORS Krangel, M.S., Band, H., Hata, S., McLean, J. and Brenner, M.B.
TITLE Structurally divergent human T cell receptor gamma proteins encoded
by distinct C gamma genes
JOURNAL Science 237 (4810), 64-67 (1987)
PUBMED 2955517
COMMENT Original source text: Human (cell line PBL Cl), cDNA to mRNA, clone
PBIC1.15.
Only a few V and J region genes occur within the genome. Diversity
in T-cell receptors occur in the V-J recombination events. In some
cases, this produces unproductively rearranged reading frames.
Also, in some recombination events, some gene regions can be
included more than one time creating even more diversity in TCR's.
FEATURES
Source
Location/Qualifiers
1..1530
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7p15-p14"
1..1530
/gene="TCRG"
<1..1530
/gene="TCRG"
/product="TCR-gamma mRNA"
153..1085
/gene="TCRG"
/note="T-cell receptor (V-J-C) precursor"
/codon_start=1
/protein_id="AA6110.1"
/db_xref="GI:339400"
/db_xref="GDB:G00-120-407"
/translation="MLSLSHASTLAVALCVAGHLEQPOISSTKTSSTKTLAEV
VSGTISATSVYWRERFGEVYQPLVSIIDGYRKSGISGKFEVDRIPEISTSL
TINAVEKODIAEYCALLEGYKFLFSGTTLVVDQLADVSPKPTIFLPSIAEK
LOKAGTYCLAEKFPDVIKIHQEKSNITLGSQBNYTKNDTYKFPSTLTPERS
LDKEHRCIVRHNNKNGVDQIIIPPIKTDVITMDPKDNCSKDANDTLQLNTSAY
YMYLLIKSVYVFAIITCCLRRTPCQNEKS"
153..197
/gene="TCRG"
/note="T-cell receptor signal peptide"
198..1082
/gene="TCRG"
/product="T-cell receptor (V-J-CI-CII-CIII)"
517..518
/gene="TCRG"
/organism="Homo sapiens"
564..565
/gene="TCRG"
/organism="Homo sapiens"
894..895
/gene="TCRG"
/organism="Homo sapiens"
942..943
/gene="TCRG"
/organism="Homo sapiens"
ORIGIN Chromosome 7p15.
Query Match 95.8%; Score 984; DB 8; Length 1530;
Best Local Similarity 98.0%; Pred. No. 4e-242;
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 12 GGGCAAAAATCAAGATTTGGTCCCGGAACAAGCTTATCATTAAGATAAACAACCT 71
DB 515 GGGAAATTATAAGAACTCTTTGGACGTGAACAACACTGTGTGCACAGATAAACAACCT 574
QY 72 TGATGACAGATGTTCCCAAGCCCAACTATTTTCTTCAATTGCTGAACAACAACCT 131
DB 575 TGAATGACAGATGTTCCCAAGCCCAACTATTTTCTTCAATTGCTGAACAACAACCT 634

QY 132 CCAAGAGGCTGGAAACATACCTTTGTCTTTGAGAAATTTTCCCTGAGATGTAATAGAT 191
DB 635 CCAAGAGGCTGGAAACATACCTTTGTCTTTGAGAAATTTTCCCTGAGATGTAATAGAT 694
QY 192 ACATGGCAAGAAAGAGAGCAACAGATTCCTGGATCCCAAGAGGAGCAACCATGAA 251
DB 695 ACATGGCAAGAAAGAGAGCAACAGATTCCTGGATCCCAAGAGGAGCAACCATGAA 754
QY 252 GACTTAAGACATACATGAAATTTAGCTGTTAACGGTCCAGAAAAGTCACTGACAA 311
DB 755 GACTTAAGACATACATGAAATTTAGCTGTTAACGGTCCAGAAAAGTCACTGACAA 814
QY 312 AGAACACAGATGTAATGCTGACATGAGAAATAATTAAGAGGTTGATCAAGAAATAT 371
DB 815 AGAACACAGATGTAATGCTGACATGAGAAATAATTAAGAGGTTGATCAAGAAATAT 874
QY 372 CTTTCCCTCCAAATAAGACGAGATGATCAATAGATCCCAAGAACAAATGTTCAAAAGA 431
DB 875 CTTTCCCTCCAAATAAGACGAGATGATCAATAGATCCCAAGAACAAATGTTCAAAAGA 934
QY 432 TGCAGATGATACATACTGCTGAGCTCACAAAACCTCTGCAATTATGATGATCCTCT 491
DB 935 TGCAATGATACATACTGCTGAGCTCACAAAACCTCTGCAATTATGATGATCCTCT 994
QY 492 CTTGCTCTCAAGATGTGTCTATTGTCATCATACCTGCTGTCTCTTAGAAGAAC 551
DB 995 CTTGCTCTCAAGATGTGTCTATTGTCATCATACCTGCTGTCTCTTAGAAGAAC 1054
QY 552 GGGTTTCTGCGAATGAGAGAAATCATTAACAGACGGTGGACCAAGAGGCCATCTTT 611
DB 1055 GGGTTTCTGCGAATGAGAGAAATCATTAACAGACGGTGGACCAAGAGGCCATCTTT 1114
QY 612 CTTGATGGTATTATGTCCTAGAAAGCTCTTGAAGATCTAGTTGGGCTTTTCTGCG 671
DB 1115 CTTGATGGTATTATGTCCTAGAAAGCTCTTGAAGATCTAGTTGGGCTTTTCTGCG 1174
QY 672 GTTTGGGCACTTTCAGTTCATGTGTGTAATTCATATTTGTAATACGGTTTTC 731
DB 1175 GTTTGGGCACTTTCAGTTCATGTGTGTAATTCATATTTGTAATACGGTTTTC 1234
QY 732 AAACCAAGTGGGACACAGAGAACCTCACTCTGTAAATGAGATGAGATGAGCCAGCA 791
DB 1235 AAACCAAGTGGGACACAGAGAACCTCACTCTGTAAATGAGATGAGATGAGCCAGCA 1294
QY 792 TCTCCAGACCAATCTCTCCATGTTTTCACAAGCTCCTCCAGCCAAACCAATAGCGCT 851
DB 1295 TCTCCAGACCAATCTCTCCATGTTTTCACAAGCTCCTCCAGCCAAACCAATAGCGCT 1354
QY 852 GCTATAGTGAACATCTCGGCTTCTAGCTTGTCCCTCTCTTATGTTCTTAAATCA 911
DB 1355 GCTATAGTGAACATCTCGGCTTCTAGCTTGTCCCTCTCTTATGTTCTTAAATCA 1414
QY 912 GATAACGCTGGAGGCTTTCATTTTACAGCCCTGAGAGGCTCTTGTGATGATGA 971
DB 1415 GATAACGCTGGAGGCTTTCATTTTACAGCCCTGAGAGGCTCTTGTGATGATGA 1474
QY 972 ATTATGATGATGTTTTCGTAATTAAGCAAAATTAATTAAGAAATTAAGAAATTT 1027
DB 1475 ATTATGATGATGTTTTCGTAATTAAGCAAAATTAATTAAGAAATTAAGAAATTT 1530
RESULT 9
BC072387 1470 bp mRNA linear PRI 30-JUN-2004
LOCUS Homo sapiens T cell receptor gamma variable 9, mRNA (CDNA clone
DEFINITION IMAGE:5587705), partial cds.
ACCESSION BC072387
VERSION BC072387.1 GI:47938917
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE
AUTHORS
Hominidae; Homo.
1 (bases 1 to 1470)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schetz T.E., Brownstein M.J., Uddin T.B., Toshiyuki S.,
Cernici P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Guarnatone P.H., Richards S.,
Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Sánchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J., Skalska U., Smailus D.E.,
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
DIRECT Submission
Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Ahler, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskett, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantiripod, S., Thomas, P.J., Touchman, J.W.,
Tougeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 174 Row: a Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORP
analysis.

FEATURES
SOURCE
Location/Qualifiers
1. 1470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5587705"
/issue_type="Ovary", pooled from 3 adults"
/clone_id="NIH_MGC_125"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1. 1470
/gene="TRGV9"
/note="synonyms: TCRGV9, VZ"
/db_xref="GeneID:6983"
/db_xref="IMGT/LIGM:TRGV9"
<1. 899

gene
CDS

ORIGIN
Query Match 94.7%; Score 973; DB 8; Length 1470;
Best Local Similarity 98.0%; Pred. No. 2.7e-239;
Matches 985; Conservatly 0; Mismatches 20; Indels 0; Gaps 0;
/gene="TRGV9"
/codon_start=3
/product="TRGV9 protein"
/protein_id="AAH72387.1"
/db_xref="GI:47938918"
/db_xref="GeneID:6983"
/db_xref="IMGT/LIGM:TRGV9"
/translation="CYGAGHLEQPOISSTKLSKARLECVVSGITISATVYMYRE
RGEVIFVSTSYDGTVRKSGIPSGKREVRIPETSTSLIHVKEKQDIAPVCA
LNEVSPMYRYKLRSGSTLVVYDKDLADVDKSPRTIPSPISAEKLDKSHKIVRIE
KFFPDVIXKHQKSKNTLGSQEPKMTKNDTYMKFSVLVPEKSLDKSHKIVRIE
MNMNGVDQRIIPPIPTDVIYMDPKNCKSKNDMDTLIDLTNSAYMYLLLLKSVV
YFALITCCILRRATPCNCKSEKS"

20 AATCAAGGATTTGGTCCCGAACAAGCTTATCAATTAAGATAAACAATTGATGACAG 79
337 ATAGAAACTCTTGGACATGGAACAACCTTTGTCAAGATTAACAACCTGATGACAG 396
80 ATGTTCCCGCAAGCCCACTATTTCTTCTTCAATTGCTGAACAAGCTCCAGAAG 139
397 ATGTTCCCGCAAGCCCACTATTTCTTCTTCAATTGCTGAACAAGCTCCAGAAG 456
140 CTGGAACATACCTTTGCTTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATACATTGGC 199
457 CTGGAACATACCTTTGCTTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATACATTGGC 516
200 AAGAAAAGAGACACACAGATCTGAGATCCGAGAGGGGAAACCATGAAGACTAAG 259
517 AAGAAAAGAGACACACAGATCTGAGATCCGAGAGGGGAAACCATGAAGACTAAG 576
260 ACAATACATGAATTTAGCTGTGTTAAGCTGACAGAAAGTCACTGACAAAGACACA 319
577 ACAATACATGAATTTAGCTGTGTTAAGCTGACAGAAAGTCACTGACAAAGACACA 636
320 GATGATCGTCAACATGAGATTAATTAAGAAAGGAGTTGATCAAGAAATTAATCTTCTC 379
637 GATGATCGTCAACATGAGATTAATTAAGAAAGGAGTTGATCAAGAAATTAATCTTCTC 696
380 CAATTAAGAGGATGATCAATCAATGATCCCAAGCAATTTGTTAAAGATGAATG 439
697 CAATTAAGAGGATGATCAATCAATGATCCCAAGCAATTTGTTAAAGATGAATG 756
440 ATACACTACTGCTGACAGCTCAACAACCTCTGATATTACATGATACCTCTGCTCC 499
757 ATACACTACTGCTGACAGCTCAACAACCTCTGATATTACATGATACCTCTGCTCC 816
500 TCAAGAGTGTCTATTATTTGATCATCATCACTGCTGTCTGTTAAGAAAGCGCTTTCT 559
817 TCAAGAGTGTCTATTATTTGATCATCATCACTGCTGTCTGTTAAGAAAGCGCTTTCT 876
560 GCTGCAATGAGAGGAATCAATTAACAGCGTGGCAAGAGGCAATTTTCTCTCATG 619
877 GCTGCAATGAGAGGAATCAATTAACAGCGTGGCAAGAGGCAATTTTCTCTCATG 936
620 GTATTGTCCTTGAAGCGCTCTTCTGAGATCTAGTTGGCTTTCTTCTGGGTTGGGC 679
937 GTATTGTCCTTGAAGCGCTCTTCTGAGATCTAGTTGGCTTTCTTCTGGGTTGGGC 996
680 CATTTAGTTCTCATGTGTGATCTATTCTATCTATTATGATTAACGTTTCAACACAGT 739
997 CATTTAGTTCTCATGTGTGATCTATTCTATCTATTATGATTAACGTTTCAACACAGT 1056
740 GGGCAGACAGAAAGCTCATCTGTAATAATGAGGAATAGCAAGGAGATCTCAGC 799
1057 GGGCAGACAGAAAGCTCATCTGTAATAATGAGGAATAGCAAGGAGATCTCAGC 1116
800 ACCAATCTCCATGTTTTCACAGGCTCTCCAGCAACCAATAGCGCTCTATAGT 859

Db 1117 ACACATCTCTGCATGTTTCCACAGCTCTCCAGCCAAACCAATAGGCGCTGCTATAGT 1176
Qy 860 GTAGACATCTGCGGCTTTCTAGCCTTGTCCTCTCTTATAGTCTTTATACAGATACTG 919
Db 1177 GTAGACATCTGCGGCTTTCTAGCCTTGTCCTCTCTTATAGTCTTTATACAGATACTG 1236
Qy 920 CCTGGAAGCCTTTTCACTTTTACACGCGCTGGAAGCAGTCTTCTTGCTAGTTGAATTAATG 979
Db 1237 CCTGGAAGCCTTTTCACTTTTACACGCGCTGGAAGCAGTCTTCTTGCTAGTTGAATTAATG 1296
Qy 980 GTGCTTTTTCGGTATATAGCAAAATTAATTTAAATAATGAAAA 1024
Db 1297 GTGCTTTTTCGGTATATAGCAAAATTAATTTAAATAATGAAA 1341
RESULT 10
HUMTCRGAAD 958 bp mRNA linear PRI 14-JUN-1995
LOCUS Homo sapiens (clone HGP02) T cell receptor gamma-chain mRNA, Cl
DEFINITION
ACCESSION M27334 GI:540463
VERSION M27334.1
KEYWORDS T cell receptor gamma-chain; constant region.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 958)
Yoshikai, Y., Toyonaga, B., Koga, Y., Kimura, N., Griessemer, H. and
Mak, T.W.
Repetitive of the human T cell gamma genes: high frequency of
nonfunctional transcripts in thymus and mature T cells
Eur. J. Immunol. 17 (1), 119-126 (1987)
2949984
JOURNAL
PUBMED
COMMENT On Sep 21, 1994 this sequence version replaced gi:341672.
Original source text: Homo sapiens cDNA to mRNA.
FEATURES
SOURCE
Location/Qualifiers
1..958
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7p15-p14"
/clone="HGP02"
/cell_type="PHA-stimulated peripheral T cell"
1..958
/gene="TCRCG1"
<1..522
/gene="TCRCG1"
/codon_start=1
/product="T cell receptor gamma chain"
/protein_id="AAA6114.1"
/db_xref="GI:540464"
/db_xref="GDB:500-120-408"
/translation="DQDLADVSPEKPTFLPSIAETKIQAKATYICLEKPPDYIKI
HMEKSNKTIIGSQDGNMTKNDITMKSMVLVPEKSLDKHRCIVRENNKNGVDQ
IIFPIKIDVITMDPKDNCSDXANDTLLOLNTSAYTYTLILKSVVYPAITICCL
LRRTAFCCNGESK"
941..946
/gene="TCRCG1"
/note="G00-120-408"
958
/gene="TCRCG1"
/note="G00-120-408"
ORIGIN
Query Match 91.3%; Score 938; DB 8; Length 958;
Best Local Similarity 99.3%; Pred. No. 2.8e-230;
Matches 953; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
Qy 61 GATAACAACCTGATGAGAGATGTTTCCCAAGCCCAACATATTTCTTCCTCAATGCT 120
Db 1 GATAACAACCTGATGAGATGTTTCCCAAGCCCAACATATTTCTTCCTCAATGCT 60

Qy 121 GAAACAAAGCTCCGAAGGCTGGAACATACCTTTGTCTTGTGAAATTTTCCCTGAT 180
Db 61 GAAACAAAGCTCCGAAGGCTGGAACATACCTTTGTCTTGTGAAATTTTCCCTGAT 120
Qy 181 GTTATTAAGATACATTTGCAAGAAAAGAGCAACAGATTTCTGGATCCCAAGAGGG 240
Db 121 GTTATTAAGATACATTTGCAAGAAAAGAGCAACAGATTTCTGGATCCCAAGAGGG 180
Qy 241 AACACATGAAGACTAAGACACATACATGAAATTTAGCTGTGTTAACGGTCCAGAAA 300
Db 181 AACACATGAAGACTAAGACACATACATGAAATTTAGCTGTGTTAACGGTCCAGAAA 240
Qy 301 TCATCGACAAAGAACACAGATTTATGCTCAGACATGAAATTAATAAAGAGTTGAT 360
Db 241 TCATCGACAAAGAACACAGATTTATGCTCAGACATGAAATTAATAAAGAGTTGAT 300
Qy 361 CAAGAAATTAATCTTCTCCCAATAAGACGGATGTCATACATATGATCCCAAGACAT 420
Db 301 CAAGAAATTAATCTTCTCCCAATAAGACGGATGTCATACATATGATCCCAAGACAT 360
Qy 421 TGTTCAAAAGATGCAAAATGATACACTAGCTGACGCTCAACAAACCTCTGATATTAC 480
Db 361 TGTTCAAAAGATGCAAAATGATACACTAGCTGACGCTCAACAAACCTCTGATATTAC 420
Qy 481 ATGTACTCTCTCTGCTCTCTCAAGATGTGTCTTATTTTGGCATCATCACTGCTGTG 540
Db 421 ACGTACTCTCTCTGCTCTCTCAAGATGTGTCTTATTTTGGCATCATCACTGCTGTG 480
Qy 541 CTGGAAGAACGGCTTCTGCTGCAATGAGAGAAATCATATACAGAGGTGGCAAGAGA 600
Db 481 CTGGAAGAACGGCTTCTGCTGCAATGAGAGAAATCATATACAGAGGTGGCAAGAGA 540
Qy 601 GGGCATCTTTTCTCATCGTTATATGTCCTTGAAGAGCTCTTGAAGATCTAGTGGC 660
Db 541 GGGCATCTTTTCTCATCGTTATATGTCCTTGAAGAGCTCTTGAAGATCTAGTGGC 600
Qy 661 TTTCTTCTGGGTTTGGGCCATTTCAAGTCTCAGTGTGTATCTTATATATTGTA 720
Db 601 TTTCTTCTGGGTTTGGGCCATTTCAAGTCTCAGTGTGTATCTTATATATTGTA 660
Qy 721 TAAAGGTTTCAACCAAGTGGGACACAGAGAACCTCCTGTATTAACAATGAGAAAT 780
Db 661 TAAAGGTTTCAACCAAGTGGGACACAGAGAACCTCCTGTATTAACAATGAGAAAT 720
Qy 781 TACGGTTTCAACCAAGTGGGACACAGAGAACCTCCTGTATTAACAATGAGAAAT 718
Db 721 TACGGTTTCAACCAAGTGGGACACAGAGAACCTCCTGTATTAACAATGAGAAAT 718
Qy 719 AGCCACGGCGATCTCCAGCAACATCTCCATGTTTCCACAGTCTCTCCAGCAACCC 778
Db 719 AGCCACGGCGATCTCCAGCAACATCTCCATGTTTCCACAGTCTCTCCAGCAACCC 719
Qy 841 AAATAGCGCTGTATAGTGTAGACATCTCGGGCTTACGCTTGTCTCTTATAGT 900
Db 779 AAATAGCGCTGTATAGTGTAGACATCTCGGGCTTACGCTTGTCTCTTATAGT 838
Qy 901 TTTCTTATCGAATACGCTGGAAGGCTTCAATTTTACAGCGCTGGAAGAGCTTCT 960
Db 839 TTTCTTATCGAATACGCTGGAAGGCTTCAATTTTACAGCGCTGGAAGAGCTTCT 898
Qy 961 TTGCTAGTTGAATATGTGTGTGTTTTCGGTATATAGCAAAATTAATTTAAAAAATG 1020
Db 899 TTGCTAGTTGAATATGTGTGTGTTTTCGGTATATAGCAAAATTAATTTAAAAAATG 958
RESULT 11
BC039725 1655 bp mRNA linear PRI 06-JUN-2005
LOCUS Homo sapiens similar to T-cell receptor gamma chain V region
DEFINITION PT-gamma-1/2 precursor, mRNA (cDNA clone MGC:47828 IMAGE:5227869),
complete cds.
ACCESSION BC039725
VERSION BC039725.1 GI:24980796
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1655)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heile, F., Datchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Abramson, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McGowan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettleman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smailus, D.E., Schmeich, A., Schein, J.E., Jones, S.J. and Marra, M.A.

CONSRM
TITLE

Mammalian Gene Collection Program Team

JOURNAL
PUBMED

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1655)

JOURNAL
PUBMED

Submitted (12-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

AUTHORS
CONSRM

NH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK
COMMENT

Contact: MGC help desk

REMARK
COMMENT

Email: cgabs-remail.nih.gov

REMARK
COMMENT

Tissue Procurement: Life Technologies, Inc.

REMARK
COMMENT

cDNA Library Preparation: Life Technologies, Inc.

REMARK
COMMENT

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

REMARK
COMMENT

Galtherburg, Maryland;

REMARK
COMMENT

Web site: <http://www.nisc.nih.gov/>

REMARK
COMMENT

Contact: nisc_mgc@hgti.nih.gov

REMARK
COMMENT

Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaas, R., Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantiridop, S., Thomas, P.J., Touchman, J.W., Turgerson, C., Vogt, J.L., Walker, W.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL ac: <http://image.lnl.gov>

REMARK
COMMENT

Series: IRAK plate: 82 Row: f Column: 20.

REMARK
COMMENT

Location/Qualifiers

FEATURES
SOURCE

1. 1655

FEATURES
SOURCE

/organism="Homo sapiens"

FEATURES
SOURCE

/mol_type="mRNA"

FEATURES
SOURCE

/db_xref="taxon:9606"

FEATURES
SOURCE

/clone="MGC:47828 IMAGE:5227869"

FEATURES
SOURCE

/tissue_type="Pancreas, Spleen, adult pooled"

FEATURES
SOURCE

/clone_id="NH_MGC_120"

FEATURES
SOURCE

/lab_host="DH10B"

FEATURES
SOURCE

/note="Vector: pCMV-SPORT6"

FEATURES
SOURCE

1. 1655

FEATURES
SOURCE

/gene="LOC442670"

FEATURES
SOURCE

/db_xref="GeneID:442670"

FEATURES
SOURCE

191. 1162

FEATURES
SOURCE

/gene="LOC442670"

FEATURES
SOURCE

/product="hypothetical protein LOC442670"

ORIGIN

/protein_id="AAH39725.1"
/db_xref="GI:24980797"
/db_xref="GeneID:442670"
/translation="MOMAAVLAIPSPASOKSNLEGRKSVIRPOTSSARTCDLA
EESNGIHWYLIHQEGAPFORLOYIYGSYNSKVLSESVSGKRYTASTNNLILRN
LLENFGVYCAWGPDPYKLFSGSTLVVDKQDAVSPRPITFLPSIABTKQK
ACTVYCLEKFPDPIIKIMOEKSNITLIGSGNMTKNDYMKRSMIVPEESLDK
EHCIVHNNKNGIDOEIIPPIKIVSTVDPKSYSDADVIMTDPKDNMSKDN
DTLLQLNTSAVYTYLLLLLSVYFAIITCCLLRATPCNGERS"

Query Match 88.0%; Score 903.8; DB 8; Length 1655;
Best Local Similarity 92.9%; Pred. No. 1,7e-221;
Matches 978; Conservative 0; Mismatches 27; Indels 48; Gaps 1;

20 AATCAAGTATTTGGTCCCGAACAAGCTTATCTACAGTAAACAAGCTGATGCG 79
552 ATAAAGAACTCTTGGACATGGAACAACCTGTGTCAAGATTAACAACCTGATGCG 611
80 ATGTTCCCGCAAGCCCACTATTTCTTCTCAATGCTGAACAAGCTCCAGAAG 139
612 ATGTTCCCGCAAGCCCACTATTTCTTCTCAATGCTGAACAAGCTCCAGAAG 671
140 CTGGAACATACCTTGTCTTCTGAGAAATTTTCCCTGATGTATTAAATATTCATGCG 199
672 CTGGAACATACCTTGTCTTCTTCTGAGAAATTTTCCCAATATTTAAGATATTCATGCG 721
200 AAGAAAAGAAAGCAACACGATTCGGATCCAGAGGGGAACACCATGAACATACG 259
732 AAGAAAAGAAAGCAACACGATTCGGATCCAGAGGGGAACACCATGAACATACG 791
260 ACACATACATGAATTAAGTGTGTTAAGGTGCAAGAAAGTCACTGGAACAAACACA 319
792 ACACATACATGAATTAAGTGTGTTAAGGTGCAAGAAAGTCACTGGAACAAACACA 861
320 GATGATCGTCAGACATGAGAAATATTAACAAAGGATGATCAAGAAATATTCCTC 379
852 GATGATCGTCAGACATGAGAAATATTAACAAAGGATGATCAAGAAATATTCCTC 911
380 CAATTAAGC-----GG 391
912 CAATTAAGCAGATGTCACACACAGATGCCAAGACATTAATTAACAAAGTCAATG 971
392 ATGTCATCAATGAATGCCAAGACATTTGTTCAAAAGTGAATGATACCTGCTGC 451
972 ATGTCATCAATGAATGCCAAGACATTTGTTCAAAAGTGAATGATACCTGCTGC 1031
452 TGCAGCTCAAAACACCTGTCATTTAATCACTGCTCCTGCTCCTCAAGAGTGTG 511
1032 TGCAGCTCAAAACACCTGTCATTTAATCACTGCTCCTGCTCCTCAAGAGTGTG 1091
512 TGTATTTTGGCATCATCACTGCTGCTGCTGTTAAGAAAGCGCTTCTGCTCAATGAG 571
1092 TGTATTTTGGCATCATCACTGCTGCTGCTGTTAAGAAAGCGCTTCTGCTCAATGAG 1151
572 AGAAATCATTAACAGAGGTGGCAACAGAGCCATCTTCCATCGGTTATTCGCC 631
1152 AGAAATCATTAACAGAGGTGGCAACAGAGCCATCTTCCATCGGTTATTCGCC 1211
632 AGAAGGCTCTTGAAGATTAAGTGGGCTTCTTCTGCGGTTTGGGCCATTCAGTTCT 691
1212 AGAAGGCTCTTGAAGATTAAGTGGGCTTCTTCTGCGGTTTGGGCCATTCAGTTCT 1271
692 CATGTGTGATCTATCTATCATTTATGATTAACGCTTTTCAACCAAGTGGGCAACAGAG 751
1272 CATGTGTGATCTATCTATCATTTATGATTAACGCTTTTCAACCAAGTGGGCAACAGAG 1331
752 AACTCACTGTGAATTAACATGAAGAACAGGCGGATCTCCAGACATCAATCTCTCC 811
1332 AACTCACTGTGAATTAACATGAAGAACAGGCGGATCTCCAGACATCAATCTCTCC 1391
812 ATGTTTCCAGACCTCTCCAGCAACCAATATAGGCGCTGATAGTGAATCATCTG 871

Db 1392 ATGTTTCCACAGCTCTCCAGCAACCAATAGCCCTGCTATAGTGTAGACAGCTG 1451
Qy 872 CGGCTCTAGCCTGTGCTCTCTCTTGTAGTCTTTAATCAGATTAACCTGGAAAGCTT 931
Db 1452 CGGCTCTAGCCTGTGCTCTCTCTTGTAGTCTTTAATCAGATTAACCTGGAAAGCTT 1511
Qy 932 TCATTTACACGCCCTGAGAGCTCTTGTGAGTGAATTAATGAGTGTGTTTCC 991
Db 1512 TCATTTACACGCCCTGAGAGCTCTTGTGAGTGAATTAATGAGTGTGTTTCC 1571
Qy 992 GTAATAGCAAAATTAATTTAAAAATGAAAA 1024
Db 1572 GTATTAAGCAAAATTAATTTAAAAATGAAAA 1604

RESULT 12
CO896252 1421 bp DNA linear PAT 05-NOV-2004
LOCUS CO896252 Sequence 76 from Patent WO2004076614.
ACCESSION CO896252
VERSION CO896252.1 GI:55468101
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Hinzmann, B., Dahl, E., Rosenenthal, A., Specht, T., Schmitz, A.,
Beckmann, G., Brummenndorf, T., Kindemann, H., Roepcke, S., Hermann, K.,
Kinzhou, L., Pilarezy, C. and Staub, E.,
Human nucleic acid sequences obtained from prostatic carcinomas
Patent: WO 2004076614-A 76 10-SEP-2004;
Hinzmann, Bernd (DB); Dahl, Edgar (DB); Rosenenthal, Andre (DB);
Specht, Thomas (DB); Schmitz, Armin (DB)
FEATURES
SOURCE 1. 1421
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 87.8%; Score 902.2; DB 6; Length 1421;
Best Local Similarity 92.8%; Pred. No. 4,5e-221;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;
Db 20 AAATCAAGTATTGTCCTCCGAAACAAAGCTTATCATTAACATTAACCTGTATGAG 79
Db 369 ATAAGAAATCTTTGGCAGTGGAAACAACATTTGTTCACAGATTAACAATTTGATGAG 428
Qy 80 ATGTTTCCCAAGCCACTATTTTCTTCTTCAATTTGCTGAACAAAGCTCCAGAAAG 139
Db 429 ATGTTTCCCAAGCCACTATTTTCTTCTTCAATTTGCTGAACAAAGCTCCAGAAAG 488
Qy 140 CTGAAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 199
Db 489 CTGAAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 548
Qy 200 AAGAAAG 259
Db 549 AAGAAAG 608
Qy 260 ACACATACATGAAATTTAGTGTGTTAACGGTGCAGAAAGTCACTGAGCAAGAACCA 319
Db 609 ACACATACATGAAATTTAGTGTGTTAACGGTGCAGAAAGTCACTGAGCAAGAACCA 668
Qy 320 GATGATGTCAGAGATGAGATTAATAAAGAGAGTGTGATCAAGAAATTAATCTTCTC 379
Db 669 GATGATGTCAGAGATGAGATTAATAAAGAGAGTGTGATCAAGAAATTAATCTTCTC 728
Qy 380 CAATTAAGAGC-----GG 391
Db 729 CAATTAAGAGAGATGTCACACAGATGATCCCAAGAGAGATTAATCAAAAGATGCAATG 788

Qy 392 ATGTGATCAAGATGATCCCAAGAGAGATTTGTTCAAAAGATGGAATGATACACTAGTC 451
Db 789 ATGTGATCAAGATGATCCCAAGAGAGATTTGTTCAAAAGATGGAATGATACACTAGTC 848
Qy 452 TGCAGCTCAAAACACCTCTGATATTTATACATGATACCTCTCTGCTCTCTCAAGGTGG 511
Db 849 TGCAGCTCAAAACACCTCTGATATTTATACATGATACCTCTCTGCTCTCTCAAGGTGG 908
Qy 512 TCTATTTGCGCATCATCAGCTGCTGTCTGTAGAAAGACGGCTTTCTGCTGCAATGAG 571
Db 909 TCTATTTGCGCATCATCAGCTGCTGTCTGTAGAAAGACGGCTTTCTGCTGCAATGAG 968
Qy 572 AGAATCATATAACAGCGTGGCAAGAGAGCCATCTTTCTCATGAGTATTTGCTCT 631
Db 969 AGAATCATATAACAGCGTGGCAAGAGAGCCATCTTTCTCATGAGTATTTGCTCT 1028
Qy 632 AGAAGCTCTTCTGAGAGATAGTGGCTTTCTTTCTGAGTTTGGGCAATTCAGTTCT 691
Db 1029 AGAAGCTCTTCTGAGAGATAGTGGCTTTCTTTCTGAGTTTGGGCAATTCAGTTCT 1088
Qy 692 CATGTGTGATCTATTTATGATTAATGATTAACGGTTTCAAAACAGTGGGACACAGAG 751
Db 1089 CATGTGTGATCTATTTATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1148
Qy 752 AACTCTCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 811
Db 1149 AACTCTCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1208
Qy 812 ATGTTTCCACAGCTCTCTCCAGCCAAACCAATAGCCCTGCTATAGTGTAGACATCTG 871
Db 1209 ATGTTTCCACAGCTCTCTCCAGCCAAACCAATAGCCCTGCTATAGTGTAGACATCTG 1268
Qy 872 CGGCTCTAGCCTGTGCTCTCTCTTGTAGTCTTTAATCAGATTAACCTGGAAAGCTT 931
Db 1269 CGGCTCTAGCCTGTGCTCTCTCTTGTAGTCTTTAATCAGATTAACCTGGAAAGCTT 1328
Qy 932 TCATTTACACGCCCTGAGAGCTCTTGTGAGTGAATTAATGAGTGTGTTTCC 991
Db 1329 TCATTTACACGCCCTGAGAGCTCTTGTGAGTGAATTAATGAGTGTGTTTCC 1388
Qy 992 GTAATAGCAAAATTAATTTAAAAATGAAAA 1024
Db 1389 GTATTAAGCAAAATTAATTTAAAAATGAAAA 1421

RESULT 13
AX332835 1421 bp DNA linear PAT 09-JAN-2002
LOCUS AX332835 Sequence 3344 from Patent WO0194629.
ACCESSION AX332835
VERSION AX332835.1 GI:18123469
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrisgan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 3344 13-DEC-2001;
JOURNAL Avalon Pharmaceuticals (US)
FEATURES
SOURCE 1. 1421
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 87.8%; Score 902.2; DB 6; Length 1421;

Best Local Similarity 92.8%; Pred. No. 4.5e-221;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

```

QY 20 AATCAAGGATTTGGTCCGGAACAAGCTTATCATTAAGATTAACAATTGATGACG 79
DB 369 ATAGAAGAACTCTTGGAGAGTGAACAACACTGTGTGTCACAGATTAACAACCTGATGACG 428
QY 80 ATGTTTCCCCCAAGCCCACTATTTTCTCTCTCAATGCTGGAACCAAGCTCCAGAGG 139
DB 429 ATGTTTCCCCCAAGCCCACTATTTTCTCTCTCAATGCTGGAACCAAGCTCCAGAGG 488
QY 140 CTGGAACATACCTTGTCTCTTGAAGAAATTTTCCCTGATGTTATTAAGATCATTTGCG 199
DB 489 CTGGAACATACCTTGTCTCTTGAAGAAATTTTCCAGATTAATTAATTAATGATTCATTGGC 548
QY 200 AAGAAAAGAGACACACAGATTTGGGATCCGAGAGGGGAAACACCATGAAGATTAACG 259
DB 549 AAGAAAAGAGACACACAGATTTGGGATCCGAGAGGGGAAACACCATGAAGATTAACG 608
QY 260 ACAATATCATGAATTTAGTGTGTTACGGTGCAGAAAGTCACTGGAACAAGAACACA 319
DB 609 ACAATATCATGAATTTAGTGTGTTACGGTGCAGAAAGTCACTGGAACAAGAACACA 668
QY 320 GATGTATCTCAGACATGAGAAATTAATAAAGGAGTTGATCAAGAAATTAATCTTCTC 379
DB 669 GATGTATCTCAGACATGAGAAATTAATAAAGGAGTTGATCAAGAAATTAATCTTCTC 728
QY 380 CAATTAAGAC-----GG 391
DB 729 CAATTAAGACAGATGTACCAACAGTGGATCCCAAGACAGTATTAATAAAGATGAATG 788
QY 392 ATGTATCATGAATTTAGTGTGTTACGGTGCAGAAAGTCACTGGAACAAGAACACA 451
DB 789 ATGTATCATGAATTTAGTGTGTTACGGTGCAGAAAGTCACTGGAACAAGAACACA 848
QY 452 TGGAGCTCAACAACACCTCTGCAATTAATTAACAATGATCTCTCTGCTCTCAAGAGTGG 511
DB 849 TGGAGCTCAACAACACCTCTGCAATTAATTAACAATGATCTCTCTGCTCTCAAGAGTGG 908
QY 512 TCTATTTGGCATCATCAGCTGTCTGTCTTGAAGAACGGCTTTCTGCTGCAATGAG 571
DB 909 TCTATTTGGCATCATCAGCTGTCTGTCTTGAAGAACGGCTTTCTGCTGCAATGAG 968
QY 572 AGAAATATTAACAAGAGTGGGACACAGAGGCCATCTTTTCTCTCATGCTGTTATGTCCT 631
DB 969 AGAAATATTAACAAGAGTGGGACACAGAGGCCATCTTTTCTCTCATGCTGTTATGTCCT 1028
QY 632 AGAAGCGTCTTGGAGATCAATGTTGGGCTTTCTTCTGAGGTTTGGGCAATTCAGTTCT 691
DB 1029 AGAAGCGTCTTGGAGATCAATGTTGGGCTTTCTTCTGAGGTTTGGGCAATTCAGTTCT 1088
QY 692 CATGTGTGATCTATTTATATTAATGATTAACGGTTTTCACACAGTGGGACACAGAG 751
DB 1089 CATGTGTGATCTATTTATATTAATGATTAATGTTTTCACACAGTGGGACACAGAG 1148
QY 752 AACCTCACTCTGATTAACAATGAGAAATGCGGCAATCTCTGACAGCAATCTCTCC 811
DB 1149 AACCTCACTCTGATTAACAATGAGAAATGCGGCAATCTCTGACAGCAATCTCTCC 1208
QY 812 ATGTTTCCACAGTCTCTCCAGCAACCAATAGCGCTGATTAAGTGAATCATCTG 871
DB 1209 ATGTTTCCACAGTCTCTCCAGCAACCAATAGCGCTGATTAAGTGAATCATCTG 1268
QY 872 CGGCTTTCAGCTTGTCTCTCTTCTTGAATTCAGATTAAGTGAATCATCTG 931
DB 1269 CGGCTTTCAGCTTGTCTCTCTCTTCTTGAATTCAGATTAAGTGAATCATCTG 1328
QY 932 TCATTTTACAGCGCTTGAAGAGCTTCTTGTCTGATTAATTAAGTGAATTCATCTG 991
DB 1329 TCATTTTACAGCGCTTGAAGAGCTTCTTGTCTGATTAATTAAGTGAATTCATCTG 1388
QY 992 GTATTAAGCAAAATTAATTAATAAATGAAAA 1024

```

```

DB 1389 GTATTAAGCAAAATTAATTAATAAATGAAAA 1421

RESULT 14
HUMTCGMH 1421 bp mRNA linear PRI 03-AUG-1993
LOCUS Human T-cell receptor aberrantly rearranged gamma-chain from cell line HPB-MLT.
DEFINITION M13231.1 GI:339168
ACCESSION M13231
VERSION 1
KEYWORDS C-region; J-region; T-cell receptor gamma chain; V-region; antigen receptor; processed gene; pseudogene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1421)
Dialynas,D.P., Murte,C., Quertermous,T., Boss,J.M., Leiden,J.M., Seidman,J.G. and Strominger,J.L.
Cloning and sequence analysis of complementary DNA encoding an aberrantly rearranged human T-cell gamma chain
Proc. Natl. Acad. Sci. U.S.A. 83 (8), 2619-2623 (1986)
3458221
JOURNAL PUBMED
Original source text: Human T-cell line HPB-MLT, cDNA to mRNA, clones PT-gamma-[1,2].
COMMENT Computer-readable sequence of [1] kindly provided by D.P. Dialynas, 31-OCT-1986.
An aberrant rearrangement at the V-J junction causes a frame shift that leads to peptide termination at the 'taa' codon at positions 370-372.
FEATURES
source
location/Qualifiers
1..1421
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
4..979
/feature="pseudo-TCG pepc"
/pseudo
/codon_start=1
4..63
/feature="pseudo-TCG signal pepc"
/pseudo
64..976
/feature="pseudo-TCG mature pepc"
/pseudo
mat_peptide
16 bp upstream of Apal site; chromosome 7.
ORIGIN
Query Match 87.8%; Score 902.2; DB 8; Length 1421;
Best Local Similarity 92.8%; Pred. No. 4.5e-221;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

```



```
Qy 493 CTGCTCTCAAGATGATGATCTAATTTTGCATCATCACTGCTGTCTGTTAGAGAACG 552
Db 481 CTGCTCTCAAGATGATGATCTAATTTTGCATCATCACTGCTGTCTGTTAGAGAACG 540
Qy 553 GCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGGACACAGAGGCCATCTTTTC 612
Db 541 GCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGGACACAGAGGCCATCTTTTC 600
Qy 613 CTGATGGGTATATGTCCTTGAACCGCTTCTGAGATCTAGTGGGCTTCTTCTGGG 672
Db 601 CTGATGGGTATATGTCCTTGAACCGCTTCTGAGATCTAGTGGGCTTCTTCTGGG 660
Qy 673 TTGGGSCATTTCACTCTCATGCTGATCTATCTATCTATCTATCTATCTATCTATCTA 732
Db 661 TTGGGSCATTTCACTCTCATGCTGATCTATCTATCTATCTATCTATCTATCTATCTA 718
Qy 733 AACCATGGGACACAGAGAACTCACTCTGTATTAACAATGAGAAATAGCCAGCGAT 792
Db 719 AACCATGGGACACAGAGAACTCACTCTGTATTAACAATGAGAAATAGCCAGCGAT 778
Qy 793 CTCAGACCAATCTCTCCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCTG 852
Db 779 CTCAGACCAATCTCTCCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCTG 838
Qy 853 CTATAGTGAACATCTCTGCGGCTTCTAGCCTTGTCCCTCTTAGTGTCTTAAATCAG 912
Db 839 CTATAGTGAACATCTCTGCGGCTTCTAGCCTTGTCCCTCTTAGTGTCTTAAATCAG 898
Qy 913 ATAACTGCTGGAAGCCTTTCATTTTAACAGCCCTGAAGCAGTCTTCTTGTAGTGA 972
Db 899 ATAACTGCTGGAAGCCTTTCATTTTAACAGCCCTGAAGCAGTCTTCTTGTAGTGA 958
Qy 973 TTATGTGTGTGTTTTCCTTAATAGCAAAATTAATTAATAAATGAAGTT 1027
Db 959 TTATGTGTGTGTTTTCCTTAATAGCAAAATTAATTAATAAATGAAGTT 1013
```

Search completed: December 10, 2005, 22:14:55
Job time : 5319 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2004, 06:11:46 : Search time 1877 Seconds
(without alignments)
1461.269 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307
Sequence: 1 MOWPSPPLFFPFLQLKQSS.....RYIGKRRRATRFMDRRGTP 58

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-O=/cgn2_1/USPRO.epool/US10031158/runat_06122004_082659_15887/app_query.fasta_1.199
-DB=genembl -QMT=faastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158 @CGN_1_1_2527 @runat_06122004_082659_15887 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenEmbl:*

1: gb_ba:*\n2: gb_hcg:*\n3: gb_in:*\n4: gb_on:*\n5: gb_ov:*\n6: gb_pat:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_ro:*\n11: gb_scs:*\n12: gb_sy:*\n13: gb_un:*\n14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	410	9	HUMTCGCD1 M14996 Human T-cell
2	307	100.0	426	6	AX884816 Sequence
3	307	100.0	426	6	BD024426 Sequence
4	307	100.0	470	6	CQ131517 Sequence

5	307	100.0	470	6	CQ290455 Sequence
6	307	100.0	477	6	CQ132293 Sequence
7	307	100.0	477	6	CQ200006 Sequence
8	307	100.0	477	6	CQ291096 Sequence
9	307	100.0	539	6	CQ409917 Sequence
10	307	100.0	539	6	CQ496769 Sequence
11	307	100.0	723	9	AJ583012 Homo sapi
12	307	100.0	726	9	AJ583014 Homo sapi
13	307	100.0	825	6	189883 Sequence 18
14	307	100.0	825	6	HSTRGSA Homo sapien
15	307	100.0	958	9	HUMTCRGAD Homo sapien
16	307	100.0	1027	6	AX074415 Sequence
17	307	100.0	1027	6	AF151103 Homo sapi
18	307	100.0	1054	9	BC062761 Homo sapi
19	307	100.0	1316	6	CQ493215 Sequence
20	307	100.0	1316	6	CQ494607 Sequence
21	307	100.0	1470	9	BC072387 Homo sapi
22	307	100.0	1530	9	HUMTCRGA Homo T-cell
23	307	100.0	1825	9	BC072396 Homo sapi
24	307	100.0	66558	2	AC130306 Homo sapi
25	307	100.0	140691	9	AF159056 Homo sapi
26	307	100.0	171816	9	AC006033 Homo sapi
27	304	99.0	400	9	HSTCELG1 Human gene
28	304	99.0	486	6	CQ493135 Sequence
29	302	98.4	316	6	CQ144550 Sequence
30	302	98.4	316	6	CQ204403 Sequence
31	302	98.4	316	6	CQ302987 Sequence
32	298	97.1	569	6	CQ511627 Sequence
33	294	95.8	330	9	HUMTCGCG Homo T-cell
34	294	95.8	635	9	HSTRGCG4 Human trunc
35	294	95.8	720	9	HSTRGCL0 Human abeyr
36	294	95.8	1013	9	HUMTCRGAA Homo sapien
37	294	95.8	1160	9	HSTRGCG4 Human mRVA
38	294	95.8	1402	9	HUMTCRGA Homo T-cell
39	294	95.8	1421	6	AX32835 Sequence
40	294	95.8	1421	9	HUMTCGXM Human T-cell
41	294	95.8	1655	9	BC039725 Homo sapi
42	292	95.1	757	9	HUMTCGCG Homo sapien
43	290	94.5	201	11	BV205008 sgmm21701
44	290	94.5	300	9	HUMTCGCE Human T-cell
45	290	94.5	330	9	HUMTCGCF Human T-cell

ALIGNMENTS

RESULT 1
HUMTCGCD1 410 bp DNA linear PRI 14-SEP-1995
LOCUS Human T-cell receptor gamma-chain (TCRGC1) gene,
DEFINITION C1-region, exon 1.
ACCESSION M14996
VERSION M14996.1 GI:339076
C-region; T cell receptor gamma-chain; T-cell receptor; germline.
SEGMENT 1 of 3
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 410)
Lefranc M.P., Forster A. and Rabbits T.H.
Genetic polymorphism and exon changes of the constant regions of
the human T-cell rearranging gene gamma
Proc. Natl. Acad. Sci. U.S.A. 83 (24), 9596-9600 (1986)
JOURNAL MEDLINE
PUBMED 87092294
COMMENT Original source text: Homo sapiens (clone: lambda-D19) DNA.
Entry revisions for [1] kindly provided by M.-P. Lefranc,
22-JUN-1988

FEATURES
Source Location/Qualifiers
1..410
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

/map="7p15-p14"
/clone="lambda-D19"
/cell_line="D-PLT"
<41..370
/gene="TCRCG1"
/note="G00-120-408"
/number=1

ORIGIN 205 bp upstream of BamHI site.

Alignment Scores:

Pred. No.:	3.57e-30	Length:	410
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x HUMTCGCD1 (1-410)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 53 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTCGTGAACAAAGCTCC 112

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 113 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 172

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 173 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAACACCA 226

RESULT 2

AX884816 426 bp DNA linear PAT 18-DEC-2003

LOCUS Sequence 679 from Patent EP1033401.

DEFINITION AX884816

VERSION AX884816.1 GI:40040355

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 Dumas Milne Edwards, J.B., Duclet, A. and Giordano, J.Y.

AUTHORS Expressed sequence tags and encoded human proteins

TITLE Patent: EP 1033401-A 679 06-SEP-2000;

JOURNAL Genset (FR)

FEATURES Location/Qualifiers

source 1..426

CDS /organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

267..>425

/note="unassigned protein product"

/codon_start=1

/protein_id="CA98930.1"

/db_xref="GI:40040356"

/translation="MKINDTYMKFSLTVBKSIDKERRCIVRHNNKNGVDQELIIP
PKTDVITM"

ORIGIN

Alignment Scores:

Pred. No.:	3.73e-30	Length:	426
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-031-158-14 (1-58) x AX884816 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

|||||

|||||

|||||

DB 94 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTCGTGAACAAAGCTCC 153

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 154 AGAAGGCTGGAAACATACCTTGTCTTCTTGAATTTTCCCTGATGTTATTAAAGATAC 213

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 214 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAACACCA 267

RESULT 3

BD024426 426 bp DNA linear PAT 27-AUG-2002

LOCUS Sequence tag and encoded human protein.

DEFINITION BD024426

ACCESION BD024426.1 GI:22565649

VERSION JP 2001269182-A/672

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 426)

AUTHORS Edwards, J.B.D.M., Duclet, E. and Jordan, J.Y.

TITLE Sequence tag and encoded human protein

JOURNAL Patent: JP 2001269182-A 672 02-OCT-2001;

GENSET

COMMENT OS Homo sapiens (human)

PN JP 2001269182-A/672

PD 02-OCT-2001

PF 24-FEB-2000 JP 2000118773

PI 26-FEB-1999 US 60/122487

PI JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES

PI JORDAN

PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC

C12N5/10,

PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N5/00, C12N5/00, PC

G06F15/40

CC Location/Qualifiers

FT 267..425.

FEATURES Location/Qualifiers

source 1..426

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

Alignment Scores:

Pred. No.:	3.73e-30	Length:	426
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-031-158-14 (1-58) x BD024426 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 94 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTCGTGAACAAAGCTCC 153

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 154 AGAAGGCTGGAAACATACCTTGTCTTCTTGAATTTTCCCTGATGTTATTAAAGATAC 213

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 214 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAACACCA 267

RESULT 4

LOCUS CQ131517 470 bp DNA linear PAT 21-JAN-2004

DEFINITION Sequence 1539 from Patent WO0157276.

ACCESSION CQ131517
VERSION CQ131517.1 GI:41088873
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 1539 09-AUG-2001;
Acemica, Inc. (US)

FEATURES
source location/Qualifiers
1..470
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN BONE MARROW, SIGNAL
= 2.2"

ORIGIN
Alignment Scores:
Pred. No.: 4.16e-30 Length: 470
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-10-031-158-14 (1-58) x CQ131517 (1-470)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 203 ATGAGAGTGTTCCTCCCAAGCCCACTATTCTTCTTCCTCAATTCGAAACAAAGCTCC 262

Qy 21 ArgArgLeuGlnIshIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 263 AGAAGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 322

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
Db 323 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATCCAGAGGGGAAACACCA 376

RESULT 5
LOCUS CQ290455 470 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 1560 from Patent WO0186003.
ACCESSION CQ290455
VERSION CQ290455.1 GI:41251032
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 1560 15-NOV-2001;
Acemica, Inc. (US)

FEATURES
source location/Qualifiers
1..470
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN LUNG, SIGNAL = 1.1"

ORIGIN
Alignment Scores:
Pred. No.: 4.16e-30 Length: 470
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ290455 (1-470)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 203 ATGAGAGTGTTCCTCCCAAGCCCACTATTCTTCTTCCTCAATTCGAAACAAAGCTCC 262

Qy 21 ArgArgLeuGlnIshIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 263 AGAAGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 322

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
Db 323 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATCCAGAGGGGAAACACCA 376

RESULT 6
LOCUS CQ132293 477 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 2315 from Patent WO0157276.
ACCESSION CQ132293
VERSION CQ132293.1 GI:41089649
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 2315 09-AUG-2001;
Acemica, Inc. (US)

FEATURES
source location/Qualifiers
1..477
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN BONE MARROW, SIGNAL
= 1.4"

ORIGIN
Alignment Scores:
Pred. No.: 4.23e-30 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-10-031-158-14 (1-58) x CQ132293 (1-477)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 212 ATGAGAGTGTTCCTCCCAAGCCCACTATTCTTCTTCCTCAATTCGAAACAAAGCTCC 271

Qy 21 ArgArgLeuGlnIshIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 272 AGAAGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 331

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
Db 332 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATCCAGAGGGGAAACACCA 385

RESULT 7
LOCUS CQ200006 477 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 2283 from Patent WO0157271.
ACCESSION CQ200006
VERSION CQ200006.1 GI:41185670
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human breast and bc 474 cells
JOURNAL Patent: WO 0157271-A 2283 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES location/Qualifiers
source 1..477
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN BT474, SIGNAL = 1.1"
ORIGIN
Alignment Scores:
Pred. No.: 4,23e-30 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-031-158-14 (1-58) x CQ200006 (1-477)
Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 212 ATGCAGATGTTTCCCCCAAGCCCACTATTCTTCTTCAATTCCTGAAACAAAGCTCC 271
Qy 21 ArgArgLeuGlnIuHsIthRphValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
Db 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAAATTTTCCCTGATGTATTAAGATAC 331
Qy 41 IleglyLysLysArgArgAlaThrArgPheTyrPaspProArgArgGlyThrPro 58
Db 332 ATTGCAAGAAAGAAAGACACGATTCCTGGATCCAGAGAGGGAACACCA 385
RESULT 8
CQ291096 477 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 2201 from Patent WO0186003.
DEFINITION CQ291096
ACCESSION CQ291096
VERSION CQ291096.1 GI:41251673
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 2201 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES location/Qualifiers
source 1..477
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN LUNG, SIGNAL = 5.3"
ORIGIN
Alignment Scores:
Pred. No.: 4,23e-30 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ291096 (1-477)
Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 212 ATGCAGATGTTTCCCCCAAGCCCACTATTCTTCTTCAATTCCTGAAACAAAGCTCC 271
Qy 21 ArgArgLeuGlnIuHsIthRphValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
Db 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAAATTTTCCCTGATGTATTAAGATAC 331
Qy 41 IleglyLysLysArgArgAlaThrArgPheTyrPaspProArgArgGlyThrPro 58
Db 332 ATTGCAAGAAAGAAAGACACGATTCCTGGATCCAGAGAGGGAACACCA 385
RESULT 9
CQ490917/c 539 bp DNA linear PAT 30-JAN-2004
LOCUS Sequence 22784 from Patent WO0160860.
DEFINITION CQ490917
ACCESSION CQ490917
VERSION CQ490917.1 GI:41456536
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 22784 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES location/Qualifiers
source 1..539
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4,84e-30 Length: 539
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-031-158-14 (1-58) x CQ490917 (1-539)
Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 410 ATGCAGATGTTTCCCCCAAGCCCACTATTCTTCTTCAATTCCTGAAACAAAGCTCC 351
Qy 21 ArgArgLeuGlnIuHsIthRphValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
Db 350 AGAAGGCTGGAACATACCTTGTCTTCTTGAAATTTTCCCTGATGTATTAAGATAC 291
Qy 41 IleglyLysLysArgArgAlaThrArgPheTyrPaspProArgArgGlyThrPro 58
Db 290 ATTGCAAGAAAGAAAGACACGATTCCTGGATCCAGAGAGGGAACACCA 237
RESULT 10
CQ496769 539 bp DNA linear PAT 30-JAN-2004
LOCUS Sequence 28636 from Patent WO0160860.
DEFINITION CQ496769
ACCESSION CQ496769
VERSION CQ496769.1 GI:41462398
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.

TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 28636 23-AUG-2001;
Millemmium Predictive Medicine, Inc. (US)
FEATURES
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4,84e-30 Length: 539
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-10-031-158-14 (1-58) x CQ496769 (1-539)
QY 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20
Db 410 ATGCAGATGTTCCCGCCAGCCCACTATTCTTCTCCATTCGTAACCAAGCTCC 351
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 350 AGAAGGCTGGAACATACCTTGCTCTTGAGAAATTTTCCCGATGTTATTAAGATAC 291
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 290 ATTGGCAGAAAGAAAGAGCAACGATTCGGATCTCCAGAGGAGCAACCA 237
RESULT 11
AJ583012 723 bp mRNA linear PRI 01-OCT-2003
LOCUS Homo sapiens partial mRNA for T-cell receptor gamma chain (TCRG
DEFINITION gene). clone 15.1.21.Bob.
ACCESSION AJ583012.1 GI:37495406
VERSION T-cell receptor gamma chain; TCRG gene.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Green,A.E., Lissina,A., Hutchinson,S.L., Temple,B., Boulter,J.M.,
Price,D.A. and Sewell,A.K.
TITLE Recognition of non-peptide antigens by human Vgamma9 Vdelta2 TCRs
requires contact with a cell of human origin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 723)
AUTHORS Green,A.E.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-2003) Green A.E., Nuffield Department of
Medicine, University of Oxford, Peter Medawar Building, South Parks
Rd, Oxford, OX1 3SY, UNITED KINGDOM
FEATURES
source
1. .723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="15.1.21.Bob"
/cell_type="gamma/delta T cell"
/rearranged
1. .723
/gene="TCRG"
1. .>723
/gene="TCRG"
/codon_start=1
/evidence=experimental
/product="T-cell receptor gamma chain"
/protein_id="CAE47524.1"

/db_xref="GI:37495407"
/translation="MAGHLEQPOISSTKTLSTKARLECVSGITISATSVYWRERPG
EVIOPLVISYDGTVRKESGISGKREVRIPETSTSLTHVHODIATVYCALMB
AKGKRIKVRGPSTKILITDKDADAVSPRTFILPSIATKTIQKGTVCLEKPP
DVTKIHWQEKSKSTFIIGSGQGNTKNDITMKSFWLTVPEESLDKEHRCIVRHNNKN
GVDOEIIFFPIKTVITMDPKDN"
4. .306
/gene="TCRG"
/note="TRGV9"
313. .363
/gene="TCRG"
/note="TRGJP"
364. .>723
/gene="TCRG"
/note="TRGCL"
ORIGIN
Alignment Scores:
Pred. No.: 6.72e-30 Length: 723
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 9
US-10-031-158-14 (1-58) x AJ583012 (1-723)
QY 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20
Db 377 ATGCAGATGTTCCCGCCAGCCCACTATTCTTCTCCATTCGTAACCAAGCTCC 436
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 437 AGAAGGCTGGAACATACCTTGCTCTTGAGAAATTTTCCCGATGTTATTAAGATAC 496
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 497 ATTGGCAGAAAGAAAGAGCAACGATTCGGATCTCCAGAGGAGCAACCA 550
RESULT 12
AJ583014 726 bp mRNA linear PRI 01-OCT-2003
LOCUS Homo sapiens partial mRNA for T-cell receptor gamma chain (TCRG
DEFINITION gene). clone 16.1.2.Wendy.
ACCESSION AJ583014.1 GI:37495410
VERSION T-cell receptor gamma chain; TCRG gene.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Green,A.E., Lissina,A., Hutchinson,S.L., Temple,B., Boulter,J.M.,
Price,D.A. and Sewell,A.K.
TITLE Recognition of non-peptide antigens by human Vgamma9 Vdelta2 TCRs
requires contact with a cell of human origin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 726)
AUTHORS Green,A.E.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-2003) Green A.E., Nuffield Department of
Medicine, University of Oxford, Peter Medawar Building, South Parks
Rd, Oxford, OX1 3SY, UNITED KINGDOM
FEATURES
source
1. .726
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="16.1.2.Wendy"
/cell_type="gamma/delta T cell"
/rearranged
1. .726
/gene="TCRG"

CDS

1. .>726
/gene="TCRG"
/codon_start=1
/product="T-cell receptor gamma chain"
/protein_id="CA47526.1"
/db_xref="GI:37495411"
/translation="MAGHLKQPOISSTKTLSTARLCEVSGITISATSVYWRERPG
EVIGFLVSIYDGVTRKESGPSGFEVDRIPESTSTLTITHNVEKODIATVYCALME
VTELGKIKIVFGPGTKLIITDKLDADVSPPRTIFLPSIATKLOKAGTYICLKEFP
PDVIRKIHMOEKSNITIGSOSGNTMKINDTYMKRSMVLVPEESLDKEHRCIVREHNK
NGVDQETIFPPIKIDVITMDPKDN"
4. .309
/gene="TCRG"
/note="TRGV9"
313. .366
/gene="TCRG"
/note="TRGJP"
367. .>726
/gene="TCRG"
/note="TRGCI"

C_region

ORIGIN

Alignment Scores:

Pred. No.: 6.79e-30 Length: 726
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-031-158-14 (1-58) x AJ583014 (1-726)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSer 20
Db 380 ATGCAGAGTGTTCCTCCCAAGCCCATATTTCTTCCTCAATCTGAAACAAAGCTCC 439

Qy 21 ArgArgLeuGlnLhiThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
Db 440 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 499

Qy 41 IleglyylvslysaArgAlaThrArgPheTrpAppProArgArgGlyThrPro 58
Db 500 ATTGGAGAAAGAAAGAGCAACAGATTCTGGATCCAGAGGAGAACCA 553

RESULT 13

189883 189883 825 bp DNA linear PAT 10-AUG-1998

LOCUS DEFINITION Sequence 18 from patent US 5723309.

ACCESSION 189883

VERSION 189883.1 GI:3409823

KEYWORDS

SOURCE . Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 825)

AUTHORS Bonneville,M.

TITLE Production of subunits of soluble T cell receptors by co-transfection

JOURNAL Patent: US 5723309-A 18-03-MAR-1998;

FEATURES

source location/Qualifiers

1. .825
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-30 Length: 825
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-031-158-14 (1-58) x 189883 (1-825)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSer 20
Db 440 ATGCAGAGTGTTCCTCCCAAGCCCATATTTCTTCCTCAATCTGAAACAAAGCTCC 499

Qy 21 ArgArgLeuGlnLhiThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
Db 500 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 559

Qy 41 IleglyylvslysaArgAlaThrArgPheTrpAppProArgArgGlyThrPro 58
Db 560 ATTGGAGAAAGAAAGAGCAACAGATTCTGGATCCAGAGGAGAACCA 613

RESULT 14

HSTCRGSA

LOCUS H. sapiens mRNA for soluble gamma TCR. 825 bp mRNA linear PRI 09-SEP-1993

DEFINITION X72500

ACCESSION X72500.1 GI:298106

VERSION soluble receptor; T cell receptor gamma chain.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Davodeau, F., Houde, I., Boulot, G., Romagne, F., Necker, A., Canavo, N., Peyrat, M.A., Hallet, M.M., Vie, H., Jacques, Y., Martuza, R. and Bonneville, M.

TITLE Secretion of disulfide-linked human T-cell receptor gamma delta heterodimers

JOURNAL J. Biol. Chem. 268 (21), 15455-15460 (1993)

MEDLINE 93340140

PUBMED 8340374

REFERENCE 2 (bases 1 to 825)

AUTHORS Davodeau, F.

TITLE Direct Submission

JOURNAL Submitted (28-APR-1993) F. Davodeau, INSERM U211, Institut de Biologie, 9 Quai Monceau, 44035 Nantes Cedex 01, FRANCE

FEATURES

source location/Qualifiers

1. .825
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q15"
/cell_type="T lymphocyte"
/tissue_type="blood"
1. .825
/codon_start=1
/product="gamma-delta T-cell receptor"
/protein_id="CA51165.1"
/db_xref="GI:298107"
/translation="MLSLHASTLVAGALCVGAGHLQPOISSTKTLSTARLECV
VSGITISATSVYWRERPGEVIGFLVSIYDGVTRKESGPSGFEVDRIPESTSTLT
TINHNVEKODIATVYCALMEAOQELGKIKIVGPGTKLIITDKLDADVSPPRTIFLPS
IAETKLOKAGTYICLKEFPDVKIRIHMEKSNITIGSOSGNTMKINDTYMKRSMVL
VPEESLDKEHRCIVREHNKGVQETIFPPIKIDVITMDPKDCKSKDANTLLIQ"

V_region

N_region

J_region

C_region

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-30 Length: 825
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB:	9	Gaps:	0
US-10-031-158-14 (1-58) x HSTCRGA (1-825)			
Qy	1	MetGlnMetPheProPheSerProPhePhePheLeuGlnLeuLeuYGIInSerSer	20
Db	440	ATGCAGATGTTTCCCCCAAGCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC	499
Qy	21	ArgArgLeuGlnHisIstHisPheValPheLeuArgAanPheSerLeuMetLeuLeuArgTyr	40
Db	500	AGAAAGCTGGAACATACCTTGTGTTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATAC	559
Qy	41	ILleglylylylyValArgArgAlaIstHisArgPheTtrPhePProAlaArgGlyTyrPro	58
Db	560	ATTGGAGAAAGAAAGAAAGAACACACGATTCGTGGATTCACAGAGGGGAAACACA	613
RESULT 15			
HUMTCRGAD			
LOCUS	HUMTCRGAD	958 bp	mRNA linear PRI 14-JAN-1995
DEFINITION	Homo sapiens (clone HGP02) T cell receptor gamma-chain mRNA, C1		
ACCESSION	M27334.1	GI:540463	
VERSION	M27334		
KEYWORDS	T cell receptor gamma-chain; constant region.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Yoshihata, Y., Toyonaga, B., Koga, Y., Kimura, N., Grieser, H. and Mak, T. W.		
TITLE	Repetitive of the human T cell gamma genes: high frequency of nonfunctional transcripts in thymus and mature T cells		
JOURNAL	Eur. J. Immunol. 17 (1), 119-126 (1987)		
MEDLINE	87133835		
PUBMED	2949984		
COMMENT	On Sep 21, 1994 this sequence version replaced gi:341672.		
FEATURES	Original source text: Homo sapiens CDNA to mRNA.		
source	location/Qualifiers		
	1..958		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/map="7p15-p14"		
	/clone="HGP02"		
	/cell_type="PHA-stimulated peripheral T cell"		
gene	1..958		
	/gene="TCRCG1"		
	<1..522		
CDs	/gene="TCRCG1"		
	/codon_start=1		
	/product="T cell receptor gamma chain"		
	/protein_id="AAA6114.1"		
	/db_xref="GI:540464"		
	/db_xref="GDB:G00-120-408"		
	/translation="DKQADADVSPKPTIFLPSIAETKLOKAGTYLCLEKFPDVVIXIKI		
	HMEKKSNTILGSGEGNTKNDITYMKESMLTPEKSLDKERCIIVRHNNKNGVDGDE		
	ILPPIKTDVITMDPKONCSADNDTLILQLTNLSAYTYTLLLLKSVVPAIITCCIL		
	LRRNRFCNGEKS"		
	941..946		
polyA_signal	/gene="TCRCG1"		
	/note="G00-120-408"		
	958		
polyA_site	/gene="TCRCG1"		
	/note="G00-120-408"		
ORIGIN			
Alignment Scores:			
Pred. No.:	9,216-30	Length:	958
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB:	9	Gaps:	0
US-10-031-158-14 (1-58) x HUMTCRGAAD (1-958)			
Qy	1	MetGlnMetPheProPsoSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer	20
Db	14	ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAAACAAGCTCC	73
Qy	21	ArgArgLeuGlnIuHsIsthRpheValPheLeuArgAsnDheSerLeuMetLeuLeuArgTyr	40
Db	74	AGAAAGCTGGAACACTACTTGTCTTCTTGTGAAGAAATTTTCCCGATGTTATTAAAGTAC	133
Qy	41	ILeGlyIuLySAArgArgAlaThArgPheTrpAspProArgArgGlyThrPro	58
Db	134	ATTGGGAAGAAAAGAGACAAACGATTCGTGGATCCAGAGGGGAACACCA	187
RESULT 16			
AX074415	AX074415	1027 bp	DNA linear PAT 06-FEB-2001
LOCUS	Sequence 13 from Patent WO0104309.		
DEFINITION	AX074415		
ACCESSION	AX074415		
VERSION	AX074415.1	GI:12710553	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Pastan,I., Bessand,M., Lee,B., Vasmatazis,G. and Wolfgang,C.		
FEATURES	T-cell receptor _g(g) alternate reading frame protein, (tarp) and		
source	uses chereot Patent: WO 0104309-A 13 18-JAN-2001;		
	UNITED STATES GOVERNMENT (US)		
	Location/Qualifiers		
	1..1027		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	74..250		
	/note="unnamed protein product; Coding region for PS-TCR		
	gamma 1 polypeptide (TARP)"		
	/codon_start=1		
	/protein_id="CAC28473.1"		
	/db_xref="GI:12710554"		
	/translation="MQMPPSPLEFFFLQLIKOSRRLEHTFVLFNFSMLRLRYGKK		
	RRATRFMDPRGRGP"		
	247..582		
	/note="unamed protein product; Coding region for PS-TCR		
	gamma 2 polypeptide (deduced amino acid sequence not		
	displayed along with DNA sequence, due to overlapping		
	CDS'g)"		
	/codon_start=1		
	/protein_id="CAC28474.1"		
	/db_xref="GI:12710555"		
	/translation="WKTNDYIVKFSWLTVEKSLDKSHRCIVRHENKNKGVDOEII		
	PIKDIATVMDPKONCSKDANDTLLQLNTSAYMYMLLLKSVVFAIITCCILRR		
	AFCNCGEKS"		
ORIGIN			
Alignment Scores:			
Pred. No.:	9,95e-30	Length:	1027
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-031-158-14 (1-58) x AX074415 (1-1027)			
Qy	1	MetGlnMetPheProPsoSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer	20
Db	74	ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAAACAAGCTCC	133

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAnpSerLeuMetLeuArgTyr 40
 Db 134 AGAAGGCTGGAACATACCTTGTCTTGTGAATAATTTTCCCTGATGTATTAAGATAC 193
 Qy 41 lIleGlyLysIysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 194 ATTGGCAAGAAAAGAGAGCAACGATTTGGGATCCAGAGGGGAAACACCA 247

RESULT 17
 AF151103 1027 bp mRNA linear PRI 03-NOV-2000
 LOCUS Homo sapiens TCRgamma alternate reading frame protein (TCRG) mRNA,
 complete cds.
 ACCESSION AF151103
 VERSION AF151103.1 GI:5758136
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Esasand,M., Vasmatazle,G., Brinkmann,U., Duray,P., Lee,B. and
 Pastan,I.
 TITLE High expression of a specific T-cell receptor gamma transcript in
 epithelial cells of the prostate
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (16), 9287-9292 (1999)
 MEDLINE 99362750
 PUBMED 10430935
 2 (bases 1 to 1027)
 Wolfgang,C.D., Esasand,M., Vincent,J.J., Lee,B. and Pastan,I.
 TARP: a nuclear protein expressed in prostate and breast cancer
 cells derived from an alternate reading frame of the T cell
 receptor gamma chain locus
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9437-9442 (2000)
 MEDLINE 20402552
 PUBMED 10931945
 3 (bases 1 to 1027)
 Esasand,M., Vasmatazle,G., Brinkmann,U., Duray,P., Lee,B. and
 Pastan,I.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-1999) Division of Basic Sciences, Laboratory of
 Molecular Biology, National Institutes of Health, National Cancer
 Institute, 37 Convent Drive, Bethesda, MD 20892, USA
 FEATURES
 source
 1..1027
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /note="transcript from an unrearranged TCRg locus"
 1..1027
 /gene="TCRG"
 1..60
 /gene="TCRG"
 /note="gamma 1.2"
 61..390
 /gene="TCRG"
 /note="gamma 1; corresponds to exon 1"
 74..250
 /gene="TCRG"
 /note="TARP"
 /codon_start=1
 /product="TCRgamma alternate reading frame protein"
 /protein_id="AAG29337.1"
 /db_xref="GI:11093514"
 /translation="MQMPFPPFLFFLLKQSRRLHPTVFNFSLMLRYIGKK
 RRATFMDPRRTTP"
 CDS
 391..438
 /gene="TCRG"
 /note="gamma 1; corresponds to exon 11"
 439..579
 /gene="TCRG"
 /note="gamma 1; corresponds to exon 111"
 C_region
 C_region
 C_region
 ORIGIN

Alignment Scores:
 Pred. No.: 9.95e-30 Length: 1027
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-031-158-14 (1-58) x AF151103 (1-1027)

Qy 1 MetGlnMetPheProProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
 Db 74 ATGCAGATGTTCCCGCAAGCCACTATTTCTTCCTTCATTCGTAACCAAGCTCC 133

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAnpSerLeuMetLeuArgTyr 40
 Db 134 AGAAGGCTGGAACATACCTTGTCTTGTGAATAATTTTCCCTGATGTATTAAGATAC 193

Qy 41 lIleGlyLysIysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 194 ATTGGCAAGAAAAGAGAGCAACGATTTGGGATCCAGAGGGGAAACACCA 247

RESULT 18
 BC062761 1054 bp mRNA linear PRI 16-DEC-2003
 LOCUS Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone
 IMAGE:4248993), partial cds.
 ACCESSION BC062761
 VERSION BC062761.1 GI:38540986
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1054)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stjepicon,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Schaeetz,T.E., Brownstein,M.J., Uddin,T.B., Tobinlyuk,I.S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gurarte,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywicki,M.I., Skalka,J., Smalls,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 2 (bases 1 to 1054)
 Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (25-NOV-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.ncl.nih.gov
 Contact: MGC help desk
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: CLONTECH
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada

REMARK
COMMENT

info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Boedert, Yaron Butterfield,
Susanna Chen, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzyzinski, Keta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Neas, Pawan Pandoh, Anna-Bles Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 51 Row: e Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria. Hexamer frequency ORF
analysis.

FEATURES
source
1. .1054
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="424893"
/issue_type="Prostate"
/clone_id="NH MGC_83"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
1. .1054
/gene="TRGV9"
/note="synonyms: TCRGV9, V2"
/db_xref="IMG/LIGM:TRGV9"
/db_xref="LocusID:6983"
1. .577
/gene="TRGV9"
/codon_start=2
/product="TRGV9 protein"
/protein_id="AA62761.1"
/db_xref="GI:38540987"
/db_xref="IMG/LIGM:TRGV9"
/db_xref="LocusID:6983"
/translation="GGQKIKVFGPGTKLIITDKOLDADVSPEKTIPLPSIAETKLOK
AGYICLLEKFPVIVIKIHQEKSNVILGQENIKVNDYKFFMLTVPEKSLDK
EHRCIVHNNKGVDEIIPPIKTDVITVDPPDNCSKDNDTLTLQLTWTSAYMY
LILLKSVVYFAITTCILRTATPCNGEKS"
74. .364
/gene="TRGV9"
/note="IGC: Region: Immunoglobulin domain constant region
subfamily"
/db_xref="CDD:cd00098"

misc_feature

ORIGIN

Alignment Scores:
Pred. No.: 1.02e-29 Length: 1054
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x BC062761 (1-1054)

OY 1 MetGlnMetPheProPserProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 69 ATGCAGATGTTCCCAAGCCCACTATTTCTTCTCAATGCTGAAACAAGCTCC 128
OY 21 ArgArgLeuGlnLysIleThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40
DB 129 AGAAGGCTGGAAACAACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 188
OY 41 TleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
DB 189 ATTGGCAAGAAAGAAAGCAACGATTCGGATCCACGAGGGGGAACACCA 242
RESULT 19

CQ493215/c
LOCUS CQ493215 1316 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 25082 from Patent WO0160860.
ACCESSION CQ493215
VERSION CQ493215.1 GI:41458834
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNALS Patent: WO 0160860-A 25082 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1. .1316
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.31e-29 Length: 1316
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ493215 (1-1316)

OY 1 MetGlnMetPheProPserProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 562 ATGCAGATGTTCCCAAGCCCACTATTTCTTCTCAATGCTGAAACAAGCTCC 503
OY 21 ArgArgLeuGlnLysIleThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40
DB 502 AGAAGGCTGGAAACAACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 443
OY 41 TleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
DB 442 ATTGGCAAGAAAGAAAGCAACGATTCGGATCCACGAGGGGGAACACCA 389
RESULT 20

CQ494607/c
LOCUS CQ494607 1316 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 26474 from Patent WO0160860.
ACCESSION CQ494607
VERSION CQ494607.1 GI:41460226
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNALS Patent: WO 0160860-A 26474 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1. .1316
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.31e-29 Length: 1316
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CG494607 (1-1316)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
 Db 562 ATGCAGATGTTCCCGCCCAAGCCCACTATTCTTCTCCCTCAATGCTGAAACAAAGCTCC 503

Qy 21 ArgArgLeuGlnLuhIstHrPheValPheLeuArgPhePheSerLeuMetLeuLeuArgTyr 40
 Db 502 AGAAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATATTTATTAAGATAC 443

Qy 41 lIleGlyLyLySArgArgAlaThrArgPheTrpAapProArgGlyThrPro 58
 Db 442 ATTGGCAAGAAAGAAAGCAACGATTCGTGGATCCAGAGGGGAAACACCA 389

RESULT 21
 BC072387 1470 bp mRNA linear PRI 30-JUN-2004
 LOCUS IMAGE:5587705, partial cds.
 DEFINITION Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone
 BC072387
 ACCESSION BC072387.1 GI:47938917
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 1 (bases 1 to 1470)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Alechul,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,J., Hong,L.,
 Stiepleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheer,T.E., Brownstein,M.J., Uedell,T.B., Toshlyak,S.,
 Carinini,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McMan,P.J.,
 McErmann,K.J., Malek,J.A., Gunatane,P.H., Richards,S.,
 Wooley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shcherenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitt,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
 Schnerich,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The J.W.A.G.S. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nih.gov
 Akhter,N., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,

Dietch,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlin,E., Kwong,P., Latic,P., Legaspi,R.,
 Maduro,O.L., Masello,C., Maekari,B., Mastrian,S.D., McCloskey,J.C.,
 McQuell,J., Pearson,R., Stancirpop,S., Thomas,P.J., Touchman,J.W.,
 Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 174 Row: a Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source
 location/Qualifiers
 1..1470
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:587705"
 /issue_type="Ovary, pooled from 3 adults"
 /clone_id="N1H MGC_125"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 <1..1470
 /gene="TRGV9"
 /note="synonym: TCRGV9, V2"
 /db_xref="IMGT/LiGM:TRGV9"
 /db_xref="LOCUSID:6983"
 <1..899
 /gene="TRGV9"
 /codon_start=3
 /product="TRGV9 protein"
 /protein_id="AAH2387.1"
 /db_xref="GI:47938918"
 /db_xref="IMGT/LiGM:TRGV9"
 /db_xref="LOCUSID:6983"
 /translation="CVYAGALDEPQISSTTSKTRLECVSGITISATSYWYRE
 LRGVIVFLVSIYDGVKESGIPSGFEVDRIPESTSLTIHNEKQDIATYCA
 LMEVQSPWYKRLKPSGTLVATKQDADVSPRTFLPSIATKLOKAGTYCLE
 KFPEDVAKIHQEKSNITLQSGNMTKTDYTKFSLMTVPKSLDKERHCIVRHE
 NKKGVQOEIIFPKITDVTIMPKDKNSKANDTLLQLNTSAYWYVLLLLKSVV
 YFAITTCILARTKFCNGERS"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,49e-29 Length: 1470
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x BC072387 (1-1470)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
 Db 391 ATGCAGATGTTCCCGCCCAAGCCCACTATTCTTCTCCCTCAATGCTGAAACAAAGCTCC 450

Qy 21 ArgArgLeuGlnLuhIstHrPheValPheLeuArgPhePheSerLeuMetLeuLeuArgTyr 40
 Db 451 AGAAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATATTTATTAAGATAC 510

Qy 41 lIleGlyLyLySArgArgAlaThrArgPheTrpAapProArgGlyThrPro 58
 Db 511 ATTGGCAAGAAAGAAAGCAACGATTCGTGGATCCAGAGGGGAAACACCA 564

RESULT 22
 HMMTCRGA 1530 bp mRNA linear PRI 14-JAN-1995
 LOCUS Human T-cell receptor gamma chain VJCI-CII region mRNA,
 DEFINITION complete cds.
 ACCESSION M16768
 VERSION M16768.1 GI:339399

KEYWORDS T-cell receptor.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1530)
AUTHORS Krangel, M.S., Band, H., Hata, S., McLean, J. and Brenner, M.B. Structurally divergent human T cell receptor gamma proteins encoded by distinct C gamma genes
TITLE Science 237 (4810), 64-67 (1987)
JOURNAL 87263363
MEDLINE 2955517
PUBMED
COMMENT Original source text: Human (cell line PBL Cl), cDNA to mRNA, clone PBLCl.15.
Only a few V and J region genes occur within the genome. Diversity in T-cell receptors occur in the V-J recombination events. In some cases, this produces unproductively rearranged reading frames. Also, in some recombination events, some gene regions can be included more than one time creating even more diversity in TCR's.

FEATURES
source location/qualifiers
1. 1530 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7p15-p14"
1. 1530 /gene="TCRG"
/gene="TCRG"
/product="TCR-gamma mRNA"
153. 1085 /gene="TCRG"
/note="T-cell receptor (V-J-C) precursor"
/codon_start=1
/protein_id="AA61110.1"
/db_xref="GI:339400"
/db_xref="GDB:G00-120-407"
/translation="MLSLHASTLAIVAGLCVYGAGLHBOPISTSTKLTARLECVVSGTTSATSVYWRERPEVIOPLVSIISIDGTAKESGIPSGFEVDRIPESTLCVTHNVEKODIATYICALLEGVYKLPFSGGTLVVTQDLADVSFKFTFLPSTAEKLDKAGTLCLEKPPVIRIKIHQEKSNLISQEQNTWKTMDTYKFMFLVPEKSLDRHCIVHNNKNGVDQELIFPIKTYITWDPDNCSDANDPILQLTWTSTAYMYLLLLKSVVYFAITTCILRTATCAENGESK"
153. 197 /gene="TCRG"
/note="T-cell receptor signal peptide"
198. 1082 /gene="TCRG"
/product="T-cell receptor (V-J-CI-CII)"
517. 518 /gene="TCRG"
/organism="Homo sapiens"
564. 565 /gene="TCRG"
/organism="Homo sapiens"
/gene="TCRG"
/organism="Homo sapiens"
894. 895 /gene="TCRG"
/organism="Homo sapiens"
942. 943 /gene="TCRG"
/organism="Homo sapiens"
ORIGIN Chromosome 7p15.

Alignment Scores:
Pred. No.: 1.55e-29 Length: 1530
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-10-031-158-14 (1-58) x HUMTCRGA (1-1530)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYglnSerSer 20
Db 577 ATGAGATGTTTCCCAAGCCCACTATTTCTTCTTCAATGCTGAAACAAAGCTCC 636
Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgPheSerLeuMetLeuLeuArgTyr 40
Db 637 AGAAGCTGGAACATACCTTGTCTTCTTGAATTTTCCCGAAGTTATTAAGATAC 696
Qy 41 ILGLYLYSLYARAGALATHRARGPheTTPAsProARgARgLYTHrPro 58
Db 697 ATTGGCAAGAAAGAGACACGATCTGGGATCCAGAGGGAGACCA 750

RESULT 23
LOCUS BC072396
DEFINITION Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone MGC:90486 IMAGE:575352), complete cds.
ACCESSION BC072396
VERSION BC072396.1 GI:47939858
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1825)
AUTHORS Strausberg, R.L., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Klauener, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.L., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheet, T.E., Brownstein, M.J., Ushed, T.B., Toohily, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulvaney, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Keteeman, M., Madan, A.C., Shvchenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, G.C., Rodriguez, A.C., Grimwood, J., Schmutz, U., Myers, R.M., Disterfeld, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marry, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1825)
Strausberg, R.
Direct Submission
Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: gdc@ncl.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@ncl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hachighi, P., Hansen, N., Ho, S.-L., Karline, E., Kong, P., Latic, P., Legaspi, R., Maduro, Q.L., Maiello, C., Maekeri, B., Macerini, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stancirip, S., Thomas, P.J., Touchman, J.W., Turgeson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL ac: <http://image.lnl.gov>
 Series: IRAX Plate: 174 Row: c Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source
 Location/Qualifiers
 1. 1825
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:90486 IMAGE:5755352"
 /issue_type="Blood, adult leukocytes"
 /clone_id="NIH MGC_118"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. 1825
 /gene="TRGV9"
 /note="synonyms: TCRGV9, V2"
 /db_xref="IMG/IMG:TRGV9"
 /db_xref="LocusID:6983"

CDS

378..1325
 /gene="TRGV9"
 /codon_start=1
 /product="TRGV9 protein"
 /protein_id="AAH72396.1"
 /db_xref="GI:47939859"
 /db_xref="IMG/IMG:TRGV9"
 /db_xref="LocusID:6983"
 /translation="MLSLHTSTLAVALGACVYAGHLEPOISSTKTLSTANLBCV
 VSGITISATVYWRERGEVIOPLVSIISYGTAKESGIPSGREVDRIETSTLT
 TIHWVKODIATYCALWEVGEIGAKIKVFGPTKLIITDKDLADVSPPTILFPLS
 IAEKLOKAGTYLILKEFPDVKIHWEEKSNTIISOGSMTKTDITMKFSMLT
 VPEKSIDENR CIVRHNNKNVDOEII FPKIKDVIITMDKDNOSKQANTLLQLT
 NTSAYTYLLILIKSVYFAITTCILRTAFCCNGERS"

ORIGIN

Alignment Scores:

Pred. No.: 1.89e-29 Length: 1825
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x BC072396 (1-1825)

OY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLeuGlnSerSer 20
 DB 817 ATGCAGATGTTCCCAAGCCCACTATTTCTTCTCAATTCTGAAACAAAGCTCC 876
 OY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyr 40
 DB 877 AGAAGGCTGGAACATACCTTGTCTCTTGAAGAAATTTTCCCTAATGTTATTAAGATAC 936
 OY 41 ILeGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
 DB 937 ATTGGGAAGAAAGAAAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 990

RESULT 24

AC130306 66558 bp DNA linear HTG 09-AUG-2002
 LOCUS AC130306 11 clone RP11-126H19 map 11, LOW-PASS

SEQUENCE SAMPLING.

AC130306
 AC130306.1 GI:22165240
 HTG; HTGS PHASE0.
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 66558)

AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORSTITLE
 JOURNAL
 COMMENT

Birtten,B., Nussbaum,C. and Lander,E.
 Homo sapiens chromosome 11, clone RP11-126H19
 Unpublished
 2 (bases 1 to 66558)
 Birtten,B., Nussbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
 Camarata,J., Cheng,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,
 Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Melidrim,J., Meneus,L., Milnova,T., Mienga,V.,
 Murphy,T., Naylor,D., Nguyen,C., Nicol,R., Nozhu,C., Notman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talmage,J.,
 Teefaye,S., Theodore,J., Topnam,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (09-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIDR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L27898
 Center clone name: 126_H_19

 * NOTE: This record contains 80 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 1 735: contig of 735 bp in length
 * 736 835: gap of 100 bp
 * 836 1572: contig of 737 bp in length
 * 1573 1672: gap of 100 bp
 * 1673 2399: contig of 727 bp in length
 * 2400 2499: gap of 100 bp
 * 2500 3220: contig of 721 bp in length
 * 3221 3320: gap of 100 bp
 * 3321 4037: contig of 717 bp in length
 * 4038 4137: gap of 100 bp
 * 4138 4881: contig of 744 bp in length
 * 4882 4981: gap of 100 bp
 * 4982 5714: contig of 733 bp in length
 * 5715 5814: gap of 100 bp
 * 5815 6541: contig of 727 bp in length
 * 6542 6641: gap of 100 bp
 * 6642 7361: contig of 720 bp in length
 * 7362 7461: gap of 100 bp
 * 7462 8214: contig of 753 bp in length
 * 8215 8314: gap of 100 bp
 * 8315 9056: contig of 742 bp in length
 * 9057 9156: gap of 100 bp
 * 9157 9916: contig of 760 bp in length
 * 9917 10016: gap of 100 bp
 * 10017 10732: contig of 716 bp in length

10733 10832: gap of 100 bp
10833 11543: contig of 711 bp in length
11544 11643: gap of 100 bp
11644 12392: contig of 749 bp in length
12393 12492: gap of 100 bp
12493 13224: contig of 732 bp in length
13225 13324: gap of 100 bp
13325 14045: contig of 721 bp in length
14046 14145: gap of 100 bp
14146 14894: contig of 749 bp in length
14895 15726: contig of 732 bp in length
15727 15826: gap of 100 bp
15827 16562: contig of 736 bp in length
16563 16662: gap of 100 bp
16663 17400: contig of 738 bp in length
17401 17500: gap of 100 bp
17501 18229: contig of 729 bp in length
18230 18329: gap of 100 bp
18330 19041: contig of 712 bp in length
19042 19141: gap of 100 bp
19142 19860: contig of 719 bp in length
19861 19960: gap of 100 bp
19961 20680: contig of 720 bp in length
20681 20780: gap of 100 bp
20781 21508: contig of 728 bp in length
21509 21608: gap of 100 bp
21609 22316: contig of 708 bp in length
22317 22416: gap of 100 bp
22417 23158: contig of 742 bp in length
23159 23258: gap of 100 bp
23259 24028: contig of 770 bp in length
24029 24128: gap of 100 bp
24129 24867: contig of 739 bp in length
24868 24967: gap of 100 bp
24968 25695: contig of 728 bp in length
25696 25795: gap of 100 bp
25796 26518: contig of 723 bp in length
26519 26618: gap of 100 bp
26619 27359: contig of 741 bp in length
27360 27459: gap of 100 bp
27460 28200: contig of 741 bp in length
28201 28300: gap of 100 bp
28301 29009: contig of 709 bp in length
29010 29109: gap of 100 bp
29110 29843: contig of 734 bp in length
29844 29943: gap of 100 bp
29944 30683: contig of 740 bp in length
30684 30783: gap of 100 bp
30784 31517: contig of 734 bp in length
31518 31617: gap of 100 bp
31618 32351: contig of 734 bp in length
32352 32451: gap of 100 bp
32452 33177: contig of 726 bp in length
33178 33277: gap of 100 bp
33278 33992: contig of 715 bp in length
33993 34092: gap of 100 bp
34093 34819: contig of 727 bp in length
34820 34919: gap of 100 bp
34920 35636: contig of 717 bp in length
35637 35736: gap of 100 bp
35737 36449: contig of 713 bp in length
36450 36549: gap of 100 bp
36550 37290: contig of 741 bp in length
37291 37390: gap of 100 bp
37391 38129: contig of 739 bp in length
38130 38229: gap of 100 bp
38230 38862: contig of 723 bp in length
38863 39062: gap of 100 bp
39063 39798: contig of 736 bp in length
39799 40608: gap of 100 bp
40609 40708: contig of 710 bp in length
40708: gap of 100 bp

40709 41454: contig of 746 bp in length
41455 41554: gap of 100 bp
41555 42312: contig of 758 bp in length
42313 42412: gap of 100 bp
42413 43160: contig of 748 bp in length
43161 43260: gap of 100 bp
43261 43996: contig of 736 bp in length
43997 44096: gap of 100 bp
44097 44825: contig of 730 bp in length
44827 44925: gap of 100 bp
44927 45661: contig of 735 bp in length
45662 45761: gap of 100 bp
45762 46504: contig of 743 bp in length
46505 46604: gap of 100 bp
46605 47333: contig of 729 bp in length
47334 47433: gap of 100 bp
47434 48167: contig of 734 bp in length
48168 48267: gap of 100 bp
48268 48994: contig of 727 bp in length
48995 49094: gap of 100 bp
49095 49842: contig of 748 bp in length
49843 49942: gap of 100 bp
49943 50670: contig of 728 bp in length
50671 50771: gap of 100 bp
50771 51509: contig of 739 bp in length
51510 51610: gap of 100 bp
51610 52357: contig of 747 bp in length
52357 52456: gap of 100 bp
52457 53183: contig of 727 bp in length
53184 53283: gap of 100 bp
53284 54003: contig of 720 bp in length
54004 54103: gap of 100 bp
54104 54830: contig of 727 bp in length
54831 54930: gap of 100 bp
54931 55685: contig of 755 bp in length
55686 55785: gap of 100 bp
55786 56514: contig of 729 bp in length
56515 56614: gap of 100 bp
56615 57347: contig of 733 bp in length
57348 57447: gap of 100 bp
57448 58165: contig of 718 bp in length
58166 58265: gap of 100 bp
58265 59004: contig of 739 bp in length

Alignment Scores:

Pred. No.: 1.05e-27 Length: 66558
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x AC130306 (1-66558)

QY 1 MetGlnPhePheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 22925 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTCAATGCTGCAACCAAGCTCC 22984
QY 21 ArgArgLeuGlnIuriIsthrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40
DB 22985 AGAAGGCTGGAACACTACTTGTCTTGTGAATTTTCCCGATGTTATTAAGATAC 23044
QY 41 IlleGlyblybAlaGArgAlaThrArgPheTrpAspProAlaGArgGlyThrPro 58
DB 23045 ATTCGCAAGAAAGAAAGCAACAGATTCGAGATCCCAAGAGGGAACACCA 23098

RESULT 25

AF159056 140691 bp DNA linear PRI 17-AUG-2001
LOCUS Homo sapiens T-cell gamma receptor locus, complete sequence.
DEFINITION AF159056
ACCESSION AF159056
VERSION AF159056.1 GI:556238
KEYWORDS

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
             Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE   1 (sites)
AUTHORS     Zhan,M., Zhou,Q., Parlee,M., McKinnell,L. and Koop,B.
TITLE       Molecular Evolution of T-cell Receptor GVI Gene Family
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 140691)
AUTHORS     Zhan,M., Wang,K., Zhou,Q., McKinnell,L., Lefranc,M., Hood,L. and
             Koop,B.
TITLE       Direct Submission
JOURNAL     Submitted (14-JUN-1999) Department of Biology, University of
             Victoria, Victoria, B.C. V8W 2Y2, Canada
             Location/Qualifiers
FEATURES
    source
        1..140691
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            complement(1060..1312)
            /rpt_family="Alu"
            repeat_region
                /rpt_type=dispersed
                2473..2643
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                4742..5030
                /rpt_family="Alu"
                /rpt_type=dispersed
            V_segment
                5892..6361
                /note="V1S1P"
                /pseudo
            repeat_region
                7626..7901
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                complement(9618..10068)
                /rpt_family="LI"
                /rpt_type=dispersed
            V_segment
                10623..11087
                /note="V1S2"
                12438..12704
                /rpt_family="Alu"
                /rpt_type=dispersed
            V_segment
                14972..15439
                /note="V1S3"
                repeat_region
                    /rpt_family="Alu"
                    /rpt_type=dispersed
                repeat_region
                    complement(16750..17026)
                    /rpt_family="Alu"
                    /rpt_type=dispersed
                repeat_region
                    complement(18763..19211)
                    /rpt_family="LI"
                    /rpt_type=dispersed
            V_segment
                19769..20237
                /note="V1S4"
                repeat_region
                    21579..21856
                    /rpt_family="Alu"
                    /rpt_type=dispersed
            V_segment
                24128..24597
                /note="V1S5"
                repeat_region
                    25909..26187
                    /rpt_family="Alu"
                    /rpt_type=dispersed
            V_segment
                28453..28922
                /note="V1S5P"
                /pseudo
            repeat_region
                30235..30510
                /rpt_family="Alu"
                /rpt_type=dispersed
            V_segment
                32787..33258
                /note="V1S6P"
                /pseudo
            repeat_region
                complement(34182..34461)
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                35546..35844
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                /rpt_family="Alu"
                /rpt_type=dispersed
            complement(36692..36979)
            /rpt_family="Alu"
            /rpt_type=dispersed
            repeat_region
                37705..37755
                /rpt_family="LTR12"
                /rpt_type=dispersed
            V_segment
                38443..38916
                /note="V1S7P"
                /pseudo
            repeat_region
                complement(40536..40675)
                /rpt_family="LI"
                /rpt_type=dispersed
            V_segment
                43134..43612
                /note="V1S8"
                repeat_region
                    48563..48819
                    /rpt_family="Alu"
                    /rpt_type=dispersed
            V_segment
                51077..51494
                /note="V5P"
                /pseudo
            repeat_region
                51803..52036
                /rpt_family="MER33"
                /rpt_type=dispersed
            repeat_region
                52772..53283
                /rpt_family="LI"
                /rpt_type=dispersed
            repeat_region
                53437..53530
                /rpt_family="Alu"
                /rpt_type=dispersed
            V_segment
                56604..56921
                /note="V2"
                complement(57552..57770)
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                58501..58778
                /rpt_family="MER7"
                /rpt_type=dispersed
            repeat_region
                complement(58904..59187)
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                65291..65456
                /rpt_family="MER20"
                /rpt_type=dispersed
            repeat_region
                67326..67600
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                complement(68123..68344)
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                72150..72241
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                complement(72324..72459)
                /rpt_family="MIR"
                /rpt_type=dispersed
            repeat_region
                72545..72825
                /rpt_family="Alu"
                /rpt_type=dispersed
            V_segment
                73785..74099
                /note="V3P"
                /pseudo
            repeat_region
                complement(75262..75557)
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                75932..76208
                /rpt_family="Alu"
                /rpt_type=dispersed
            repeat_region
                complement(77446..77736)
                /rpt_family="Alu"
                /rpt_type=dispersed
            V_segment
                77681..78190
                /rpt_family="Alu"
                /rpt_type=dispersed

```

	/note="V6P"	
	/pseudo	
repeat_region	complement(78838,.79000)	
	/ <i>rpt_family</i> = <i>"MIR"</i>	
	/ <i>rpt_type</i> =dispersed	
repeat_region	complement(80629,.81096)	
	/ <i>rpt_family</i> = <i>"U1"</i>	
	/ <i>rpt_type</i> =dispersed	
V_segment	81990..82202	
	/note= <i>"Y4P"</i>	
	/pseudo	
repeat_region	complement(82530,.83651)	
	/ <i>rpt_family</i> = <i>"LTR1"</i>	
	/ <i>rpt_type</i> =dispersed	
repeat_region	complement(82876,.83159)	
	/ <i>rpt_family</i> = <i>"Alu"</i>	
	/ <i>rpt_type</i> =dispersed	
repeat_region	85826..86088	
	/ <i>rpt_family</i> = <i>"MER2"</i>	
	/ <i>rpt_type</i> =dispersed	
repeat_region	86291..86368	
	/ <i>rpt_type</i> =dispersed	
repeat_region	86542..86631	
	/ <i>rpt_family</i> = <i>"MER5"</i>	
	/ <i>rpt_type</i> =dispersed	
repeat_region	complement(90148,.90424)	
	/ <i>rpt_family</i> = <i>"Alu"</i>	
	/ <i>rpt_type</i> =dispersed	
J_segment	97655..97996	
	/note= <i>"JPI"</i>	
	/pseudo	
repeat_region	complement(99475,.99769)	
	/ <i>rpt_family</i> = <i>"Alu"</i>	
	/ <i>rpt_type</i> =dispersed	
J_segment	100267..100327	
	/note= <i>"JP"</i>	
	/pseudo	
repeat_region	complement(103808,.104066)	
	/ <i>rpt_family</i> = <i>"Alu"</i>	
	/ <i>rpt_type</i> =dispersed	
J_segment	104364..104423	
	/note= <i>"JI"</i>	
	/pseudo	
repeat_region	107283..107592	
	/ <i>rpt_family</i> = <i>"Alu"</i>	
	/ <i>rpt_type</i> =dispersed	
repeat_region	108009..108085	
	/ <i>rpt_family</i> = <i>"MR"</i>	
	/ <i>rpt_type</i> =dispersed	
C_region	108237..108566	
	/note= <i>"C1 exon 1"</i>	
Alignment Scores:		
Pred. No.:	2_42e-27	length: 140691
Scores:	307.00	Matches: 58
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0
<hr/>		
US-10-031-158-14 (1-58) x AFI159056 (1-140691)		
Oy	1 MetGlnMetpheProSerProLeuPhenheleuGlneuleuygginserSet	20
Dd	108249 ATGCAGATGGTTCCGCCAAGCCCACATTCTTTCCTCAATGCTGAACAACAACTCC	108308
Oy	21 ArGaAgLeuGIuhIstRheVaIPhelEuArGenPheserleuNeuletueuaGTyr	40
Dd	108309 AGAAGCGTGGAACATAACCCTTGTCTTCGTGAGAATTTTTTCCCTGATGTTATTAAAGATAC	108368
Oy	41 IleglYLvLysArgArgalathrAtarghetPDaspProArgrglYthrPr	58
Dd	108369 ATTGGCAAGAAAAGAAAGAACACCATCTTGGAATCCCAAGAGGGAAACCCA	108422
RESULT 26		

LOCUS	AC006033	171816 bp	DNA	linear	PRI 31-JAN-2004
DEFINITION	Homo sapiens BAC clone Rp11-121a8 from 7, complete sequence.				
ACCESSION	AC006033				
VERSION	AC006033.2 GI:4309948				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 171816) Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H., Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R., Wyllie K., Seaton M., Becker M.C., O'Laughlin M.D., Schaller M.E., Fowell G.A., Delahunty K.D., Miner T.L., Nash W.E., Cordes M., Du H., Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Jask A., Vahbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J., Ozerky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A., Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S., Kozlowicz-Reilly A., Leonard S., Kohling T., Rock S.M., Tin-Wollam A.M., Abbott A., Minx P., Maun J., Stromatt C., Latreille P., Miller N., Johnson D., Murray J., Moesener J.P., Wendl M.C., Yang S.P., Schultz B.R., Wallis J.M., Spieth J., Bieri T.A., Nelson J.O., Berkowicz N., Wohlmann P.E., Cook L.L., Hickenbotham M.T., Eldred J., Williams D., Bedell D.A., Mardis E.R., Gillett S.W., Chissoe S.L., Maitra M.A., Raymond C., Haugen E., Gillett M., Zhou Y., James R., Phelps K., Iadonoto S., Furey K., Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Pugh T.S., Baerbach R.A., Brent M.R., Keibler E., Flicker P., Bork P., Suyama M., Bailey J.A., Portnoy M.E., Torrente D., Chinwalla A.T., Ghosh W.R., Eddy S.R., McPherson J.D., Olson M.V., Eichler E.A., Green E.D., Waterston R.H. and Wilson R.K.				
TITLE	The DNA sequence of human chromosome 7				
JOURNAL	Nature 424 (6945), 157-164 (2003)				
MEDLINE	22737999				
PUBMED	12853948				
REFERENCE	2 (bases 1 to 171816)				
AUTHORS	Arnett C., Le T.P. and Wohlmann P.				
TITLE	The sequence of Homo sapiens BAC clone Rp11-121a8				
JOURNAL	Unpublished (2001)				
REFERENCE	3 (bases 1 to 171816)				
AUTHORS	Waterston R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-NOV-1998) Genome Sequencing Center, Washington				
REFERENCE	Submitted (22-NOV-1998) Genome Sequencing Center, Washington				
AUTHORS	MO 63108, USA				
TITLE	4 (bases 1 to 171816)				
AUTHORS	Waterston R.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-FEB-1999) Department of Genetics, Washington				
REFERENCE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
AUTHORS	5 (bases 1 to 171816)				
TITLE	Waterston R.				
JOURNAL	Submitted (21-DEC-1999) Department of Genetics, Washington				
REFERENCE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
AUTHORS	6 (bases 1 to 171816)				
TITLE	Waterston R.				
JOURNAL	Submitted (29-APR-2003) Department of Genetics, Washington				
REFERENCE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
AUTHORS	7 (bases 1 to 171816)				
TITLE	Wilson R.				
JOURNAL	Submitted (31-JUN-2004) Department of Genetics, Washington				
REFERENCE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
AUTHORS	On Mar 1, 1999 this sequence version replaced gi:3907471.				
TITLE	----- Genome Center				
JOURNAL	Center: Washington University Genome Sequencing Center				
COMMENT	Center code: WUGSC				
	Web site: http://genome.wustl.edu				
	Contact: sapiens@wustl.wustl.edu				
	----- Summary Statistics				

Center project name: H_NH0121A08

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPc1-11 human BAC library was made from the blood of one male donor, as described by Ooegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Catanesse, J. V., and de Jong, P. V. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-121A8
Actual end is at 171816 of RP11-121A8.

Location/Qualifiers

FEATURES

SOURCE

```
1..171816
  /organism="Homo sapiens"
  /mol_type="Genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="7"
  /map="7"
  /clone="RP11-121A8"
  /clone_11b="RPc1-11"
  446..928
    repeat_region
      /rpt_family="L2"
      1675..1829
        repeat_region
          /rpt_family="MIR"
          2729..2887
            repeat_region
              /rpt_family="L2"
              3315..3420
                repeat_region
                  /rpt_family="MIR"
                  3941..3976
                    repeat_region
                      /rpt_family="AT_rich"
                      4114..4171
                        repeat_region
                          /rpt_family="MIR"
                          4175..4305
                            repeat_region
                              /rpt_family="MIR1_type"
                              4623..4781
                                repeat_region
                                  /rpt_family="MIR2_type"
                                  4889..5191
                                    repeat_region
                                      /rpt_family="L2"
                                      5209..5499
                                        repeat_region
                                          /rpt_family="MALR"
                                          5512..5623
                                            repeat_region
                                              /rpt_family="(TA)n"
                                              7420..7746
                                                repeat_region
                                                  /rpt_family="MIR2_type"
                                                  8185..8258
                                                    repeat_region
```

```
repeat_region
  /rpt_family="L2"
  8462..9500
    repeat_region
      /rpt_family="L2"
      10020..10040
        repeat_region
          /rpt_family="(TG)n"
          10941..11048
            repeat_region
              /rpt_family="MIR2_type"
              11050..11275
                repeat_region
                  /rpt_family="MIR2_type"
                  11304..11380
                    repeat_region
                      /rpt_family="(TATG)n"
                      11436..11538
                        repeat_region
                          /rpt_family="L1"
                          11586..12300
                            repeat_region
                              /rpt_family="L1"
                              12313..13051
                                repeat_region
                                  /rpt_family="L1"
                                  14179..14497
                                    repeat_region
                                      /rpt_family="Alu"
                                      14583..14644
                                        repeat_region
                                          /rpt_family="MIR"
                                          15408..15437
                                            repeat_region
                                              /rpt_family="AT_rich"
                                              15974..16255
                                                repeat_region
                                                  /rpt_family="Alu"
                                                  17692..17733
                                                    repeat_region
                                                      /rpt_family="AT_rich"
                                                      18033..18147
                                                        repeat_region
                                                          /rpt_family="(TTTC)n"
                                                          18148..18401
                                                            repeat_region
                                                              /rpt_family="Alu"
                                                              19188..19418
                                                                repeat_region
                                                                  /rpt_family="MIR1_type"
                                                                  19779..19908
                                                                    repeat_region
                                                                      /rpt_family="MALR"
                                                                      20266..20321
                                                                        repeat_region
                                                                          /rpt_family="L2"
                                                                          22450..22481
                                                                            repeat_region
                                                                              /rpt_family="AT_rich"
                                                                              22579..22887
                                                                                repeat_region
                                                                                  /rpt_family="Alu"
                                                                                  23438..23475
                                                                                    repeat_region
                                                                                      /rpt_family="AT_rich"
                                                                                      23476..23944
                                                                                        repeat_region
                                                                                          /rpt_family="L1"
                                                                                          26364..26491
                                                                                            repeat_region
                                                                                              /rpt_family="L1"
                                                                                              26619..26798
                                                                                                repeat_region
                                                                                                  /rpt_family="L1"
                                                                                                  27125..27240
                                                                                                    repeat_region
                                                                                                      /rpt_family="MALR"
                                                                                                      27412..27433
                                                                                                        repeat_region
                                                                                                          /rpt_family="AT_rich"
                                                                                                          27672..27885
                                                                                                            repeat_region
                                                                                                              /rpt_family="Alu"
```

Alignment Scores:

```
pred. No.: 3.02e-27      length: 171816
score: 307.00           matches: 58
percent similarity: 100.00% conservative: 0
best local similarity: 100.00% mismatches: 0
query match: 100.00%     indels: 0
db: 9                   gaps: 0
```

US-10-031-158-14 (1-58) x AC006033 (1-171816)

```
Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLeuGlnSerSer 20
Db 86565 ATGCAGATGTTCCCAAGCCCACTATTTCTTCCTTCATTCGTAACAAAGCTCC 86506
Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgGlnPheSerLeuMetLeuLeuArgTyr 40
Db 86505 AGAAGCTGGAAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 86446
```

Qy 41 11GG1YLYSLYSARGALAThArgPheTrpAspProArgArgLYThPro 58
 Db 86445 ATTGGCAAGAAAGAAAGACACGATTCGGATTCACAGAGGGAACACCA 86392

RESULT 27
 HSTCELG1 400 bp DNA linear PRI 02-DEC-2003
 LOCUS Human gene fragment for T-cell receptor gamma 1 chain constant
 DEFINITION region (loc. on chromosome 7p15).
 ACCESSION X02766
 VERSION X02766.1 GI:36779
 KEYWORDS T-cell receptor.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Murre, C., Maidmann, R.A., Morton, C.C., Bongiovanni, K.F.,
 Waldmann, T.A., Shows, T.B. and Seidman, J.G.
 Human gamma-chain genes are rearranged in leukaemic T cells and map
 to the short arm of chromosome 7
 Nature 316 (6028), 549-552 (1985)

JOURNAL
 MEDLINE 85296267
 PUBMED 3875797

FEATURES
 source
 1..400
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..27
 /note="put. intron"
 <28..357
 /note="unnamed protein product; (91) C region fragment
 (357 is 1st base in codon)"
 /codon_start=3
 /protein_id="CAE82035.1"
 /db_xref="GI:38635935"
 /translation="KQIDADVSPPRTFLPSIAETKIQAKTYICLAEKPPDYIKIH
 WEERKSNITLGSQSGNTWKNTDTYMKFSWLTVPKSLDKERCIVRHNNKNVDQEI
 IFPIPKT"
 358..>400
 /note="put. intron"
 358..>400

intron
 intron

ORIGIN
 Alignment Scores:
 Pred. No.: 8.44e-30 Length: 400
 Score: 304.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.28% Mismatches: 0
 Query Match: 99.02% Indels: 0
 Gaps: 0

US-10-031-158-14 (1-58) x HSTCELG1 (1-400)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuLYGlnSerSer 20
 Db 40 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCAATGCTGAAACAAACCTCC 99

Qy 21 ArgArgLeuGlnLuh1stThPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTYr 40
 Db 100 AGAAGGCTGGAACATCCTTGTCTCTTGGAATTTTTCCTCATATTATTAAGATAC 159

Qy 41 11GG1YLYSLYSARGALAThArgPheTrpAspProArgArgLYThPro 58
 Db 160 ATTGGCAAGAAAGAAAGACACGATTCGGATTCACAGAGGGAACACCA 213

RESULT 28
 CQ493135/c 486 bp DNA linear PAT 30-JAN-2004
 LOCUS Sequence 25002 from Patent WO0160860.
 DEFINITION CQ493135
 ACCESSION CQ493135
 VERSION CQ493135.1 GI:41458754

KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
 TITLE Genes differentially expressed in human prostate cancer and their
 use
 JOURNAL Patent: WO 0160860-A 25002 23-AUG-2001;
 Millennium Predictive Medicine, Inc. (US)
 FEATURES
 source
 1..486
 Location/Qualifiers
 1..486
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.05e-29 Length: 486
 Score: 304.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.28% Mismatches: 0
 Query Match: 99.02% Indels: 0
 Gaps: 0

US-10-031-158-14 (1-58) x CQ493135 (1-486)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuLYGlnSerSer 20
 Db 411 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCAATGCTGAAACAAACCTCC 352

Qy 21 ArgArgLeuGlnLuh1stThPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTYr 40
 Db 351 AGAAGGCTGGAACATCCTTGTCTCTTGGAATTTTTCCTCATATTATTAAGATAC 292

Qy 41 11GG1YLYSLYSARGALAThArgPheTrpAspProArgArgLYThPro 58
 Db 291 ATTGGCAAGAAAGAAAGACACGATTCGGATTCACAGAGGGAACACCA 238

RESULT 29
 CQ144550 316 bp DNA linear PAT 21-JAN-2004
 LOCUS Sequence 14572 from Patent WO0157276.
 ACCESSION CQ144550
 VERSION CQ144550.1 GI:41101922
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
 TITLE Human genome-derived single exon nucleic acid probes useful for
 analysis of gene expression in human bone marrow
 JOURNAL Patent: WO 0157276-A 14572 03-AUG-2001;
 Aeomica, Inc. (US)
 FEATURES
 source
 1..316
 Location/Qualifiers
 1..316
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="MAP TO AF159056.1-EXPRESSED IN BONE MARROW, SIGNAL
 = 2.2-SWISSPROT HIT: P03986, EVALU6 7.00e-58-NT HIT:
 M14996.1, EVALU6 0.00e+00-EST_HUMAN HIT: BF679123.1,
 EVALU6 0.00e+00"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.17e-29 Length: 316
 Score: 302.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 98.37% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ204403 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYegInSerSerArg 21
DB 2 CAGATGTTCCCAAGCCCACTATTCTTCCCTCAATTCGTAACAAAGCTCCAGA 61

QY 22 ArgLeuGlnIsthrPheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyrIle 41
DB 62 AGGCTGGAACATACCTTGTCTTCCTTGAGAAATTTTCCCGATGTTATTAGATACATT 121

QY 42 GlyLysLysArgArgAlaThrArgPheTrrPaspProArgArgIYThrPro 58
DB 122 GCGAAGAAAGAGAGCAACGATTCTGGATCCAGAGGAGCAACCA 172

RESULT 30
LOCUS CQ204403 316 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 6680 from Patent WO0157271.
ACCESSION CQ204403
VERSION CQ204403.1 GI:41190250
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human breast and bc 474 cells
JOURNAL Patent: WO 0157271-A 6680 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source location/Qualifiers
1..316
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN BT474, SIGNAL = 1.1-SWISSPROT HIT: P03986, EVALUE 7.00e-58-NT HIT: M14996.1, EVALUE 0.00e+00-EST_HUMAN HIT: Bf679123.1, EVALUE 0.00e+00"

ORIGIN
Alignment Scores:
Pred. No.: 1.17e-29 Length: 316
Score: 302.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.37% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ204403 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYegInSerSerArg 21
DB 2 CAGATGTTCCCAAGCCCACTATTCTTCCCTCAATTCGTAACAAAGCTCCAGA 61

QY 22 ArgLeuGlnIsthrPheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyrIle 41
DB 62 AGGCTGGAACATACCTTGTCTTCCTTGAGAAATTTTCCCGATGTTATTAGATACATT 121

QY 42 GlyLysLysArgArgAlaThrArgPheTrrPaspProArgArgIYThrPro 58
DB 122 GCGAAGAAAGAGAGCAACGATTCTGGATCCAGAGGAGCAACCA 172

RESULT 31
LOCUS CQ302987 316 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 14092 from Patent WO0186003.
ACCESSION CQ302987
VERSION CQ302987.1 GI:41263564

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 14092 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES
source location/Qualifiers
1..316
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AF159056.1-EXPRESSED IN LUNG, SIGNAL = 1.1-SWISSPROT HIT: P03986, EVALUE 7.00e-58-NT HIT: M14996.1, EVALUE 0.00e+00-EST_HUMAN HIT: Bf679123.1, EVALUE 0.00e+00"

ORIGIN
Alignment Scores:
Pred. No.: 1.17e-29 Length: 316
Score: 302.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.37% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ302987 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYegInSerSerArg 21
DB 2 CAGATGTTCCCAAGCCCACTATTCTTCCCTCAATTCGTAACAAAGCTCCAGA 61

QY 22 ArgLeuGlnIsthrPheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyrIle 41
DB 62 AGGCTGGAACATACCTTGTCTTCCTTGAGAAATTTTCCCGATGTTATTAGATACATT 121

QY 42 GlyLysLysArgArgAlaThrArgPheTrrPaspProArgArgIYThrPro 58
DB 122 GCGAAGAAAGAGAGCAACGATTCTGGATCCAGAGGAGCAACCA 172

RESULT 32
LOCUS CQ511627/c 569 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 43494 from Patent WO0160860.
ACCESSION CQ511627
VERSION CQ511627.1 GI:41477891
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 43494 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source location/Qualifiers
1..569
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 7.4e-29 Length: 569
Score: 298.00 Matches: 57
Percent Similarity: 98.28% Conservative: 0
Best Local Similarity: 98.28% Mismatches: 1

Query Match: 97.07% Indels: 0
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ511627 (1-569)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuYsGlnSerSer 20
 Db 460 ATGCAGATGTTTCCCCCAGGCCCACTATTTTCTTCAATTGCTGAACAAAGCTCC 401

Qy 21 ArgArgLeuGlnIuHsrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
 Db 400 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCGATGATATTAAGATGC 341

Qy 41 ILeGlyLysLysAArgAArgAlaThrArgPheThrPheProAArgGlyThrPro 58
 Db 340 ATTGCGAAGAAAGAGAGACAGATTCGGATCCAGAGGGGGAACACCA 287

RESULT 33
 HUMTCGCG

LOCUS HUMTCGCG 330 bp DNA linear PRI 13-JAN-1995
 DEFINITION Human T-cell receptor gamma-chain (TCRGC2) gene C2-region
 allele gamma-2b, exon 1.
 ACCESSION M15002
 VERSION M15002.1 GI:339083
 KEYWORDS C-region; T cell receptor gamma-chain; T-cell receptor; germline.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 330)
 LeFranc,M.P., Forster,A. and Rabbits,T.H.
 Genetic polymorphism and exon changes of the constant regions of
 the human T-cell rearranging gene gamma
 Proc. Natl. Acad. Sci. U.S.A. 83 (24), 9596-9600 (1986)
 JOURNAL MEDLINE 87092294
 PUBMED 2879283

COMMENT Original source text: Human B-cell polyclonocytic leukemia line
 D-PL DNA, clone lambda-D7.
 Entry revisions for (1) kindly provided by M.-P. LeFranc,
 22-JUN-1988.

FEATURES
 source location/Qualifiers
 1..330
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="7p15-p14"
 1..330
 /gene="TCRGC2"
 1..330
 /gene="TCRGC2"
 /note="G00-120-409"
 /number=1

ORIGIN 165 bp upstream of BamHI site.

Alignment Scores:
 Pred. No.: 1.32e-28 Length: 330
 Score: 294.00 Matches: 55
 Percent Similarity: 98.28% Conservative: 2
 Best Local Similarity: 94.83% Mismatches: 1
 Query Match: 95.77% Indels: 0
 DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HUMTCGCG (1-330)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuYsGlnSerSer 20
 Db 13 ATGCAGATGTTTCCCCCAGGCCCACTATTTTCTTCAATTGCTGAACAAAGCTCC 72

Qy 21 ArgArgLeuGlnIuHsrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
 Db 73 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCGATGATATTAAGATGC 132

Qy 41 ILeGlyLysLysAArgAArgAlaThrArgPheThrPheProAArgGlyThrPro 58

Db 133 ATTGCGAAGAAAGAGAGACACGATTCGGATCCAGAGGGGGAACACCA 186

RESULT 34
 HSTRGC54
 LOCUS HSTRGC54 635 bp mRNA linear PRI 20-MAR-1991
 DEFINITION Human truncated mRNA from TRG gamma gene V(g)8-J(g)P2-C(g)2 with
 nonproductive V-J arrangement.
 ACCESSION X06776
 VERSION X06776.1 GI:37343
 KEYWORDS constant region; joining region; T-cell receptor; T-cell receptor
 gamma; variable region.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 635)
 Tighe,L., Forster,A., Clark,D.M., Boylston,A.W., Lavenit,I. and
 Rabbits,T.H.
 Unusual forms of T cell gamma mRNA in a human T cell leukemia cell
 line: implications for gamma gene expression
 Eur. J. Immunol. 17 (12), 1729-1736 (1987)
 JOURNAL MEDLINE 88083067
 PUBMED 2961573

COMMENT pM17c54 is derived from TRG gamma locus allele 2;
 see X06775 for further allele 2 nonproductive cDNA; see X06774 for
 MOLT17 TRG gamma allele 1 derived productive cDNA.
 Location/Qualifiers
 1..635
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="pM17c54"
 /cell_line="MOLT17"
 /clone_1lb="lambda gt10"
 2..217
 /note="unnamed protein product; Protein sequence is in
 conflict with the conceptual translation; V(g)8 region (AA
 1-72)"
 /codon_start=1
 /protein_id="CAA29944.1"
 /db_xref="GI:1335357"
 /translation="LPVENAVYTHWYTHOEGKAPQRLYYDYSNRYVLESISREKY
 HTYASTGKSLKFLIENLIERDSEGYICA"
 218..227
 /note="out of frame joint V(g)8-J(g)P2"
 219..635
 /note="unnamed protein product; Protein sequence is in
 conflict with the conceptual translation; pot. aberrant
 translation product"
 /codon_start=1
 /protein_id="CAA29945.1"
 /db_xref="GI:1335358"
 /translation="PWRWIKTPAKGRLIVTSPDKOLDADVSPKRTFLPSIATKLO
 KAGTYICLEKRPFDIYIKIHQEKSTLIGSEGNMKNDYMKESWMLVPEESLD
 KERRCIYRHENNKNGIDQETIFPPIYTMCLHIKKK"
 228..276
 /note="J(g)P2 region (truncated at 5'end)"
 277..605
 /note="C(g)2 (exon 1)"
 558..563
 /note="pot. polyA signal"
 596..601
 /note="pot. polyA signal"
 606..635
 /note="translated intron + polyA sequence"

ORIGIN misc_feature
 misc_feature
 misc_feature
 misc_feature

Alignment Scores:
 Pred. No.: 2.73e-28 Length: 635
 Score: 294.00 Matches: 55
 Percent Similarity: 98.28% Conservative: 2
 Best Local Similarity: 94.83% Mismatches: 1


```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pm17c64"
```

REFERENCE 1 (bases 1 to 1402)

5.36e-28	Length:	1160
----------	---------	------

TITLE
 Structurally divergent human T cell receptor gamma proteins encoded by distinct C gamma genes
JOURNAL
 Science 237 (4810), 64-67 (1987)
MEDLINE
 8726353
PUBMED
 2955517
COMMENT
 Original source text: Human (cell line IDP2), cDNA to mRNA, clone IDP2.11.
 Only a few V and J region genes occur within the genome. Diversity in T-cell receptors occur in the V-J recombination events. In some cases, this produces unproductively rearranged reading frames. Also, in some recombination events, some gene regions can be included more than one time creating even more diversity in TCR's.

FEATURES
 source
 1..1402
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="7p15-p14"
 1..1402
 /gene="TCRG"
 <1..1402
 /gene="TCRG"
 /product="TCR-gamma mRNA"
 <1..957
 /gene="TCRG"
 /note="T-cell receptor (V-J-C) precursor"
 /codon_start=1
 /protein_id="AAA61124.1"
 /db_xref="GI:339409"
 /db_xref="GDB:G00-120-407"
 /translation="GGPRTSSKTLTKSLKARLECVSGITTSATSVVWYRRPGEVIOF
 LVSSIDYCTVAKESGIPSGKEVDRIPTSTSTLTINHEKDIATYCALMEGYKK
 LFSGITTLVTDKQLDADVSFPRTFLPSIAETLQKAGTLCLEKFPDIKIHW
 EKSNGTILSQSGEIMKNDNTYMEKSMULTVEKSLDKSHRCIVRHNKNKIDIEL
 PPIITDVTVDPKDSYKQANDVTVDKRYVSKANDVITMDPKNCSKANDTLTL
 QLTNSAYVTLLLLKSVVFAITCLLRATFCNGKRS"
 293..294
 /gene="TCRG"
 /organism="Homo sapiens"
 340..341
 /gene="TCRG"
 /organism="Homo sapiens"
 670..671
 /gene="TCRG"
 /organism="Homo sapiens"
 716..717
 /gene="TCRG"
 /organism="Homo sapiens"
 766..767
 /gene="TCRG"
 /organism="Homo sapiens"
 814..815
 /gene="TCRG"
 /organism="Homo sapiens"
 7p15.

ORIGIN
 Chromosome 7p15.

Alignment Scores:
 Pred. No.: 6,628-28 Length: 1402
 Score: 294.00 Matches: 55
 Percent Similarity: 98.28% Conservative: 2
 Best Local Similarity: 94.83% Mismatches: 1
 Query Match: 95.77% Indels: 0
 DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HUMTCRGB (1-1402)

QY 1 MetGlnMetPheProProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20
 Db 353 ATGCAGATGTTTCCCCCAAGCCCACTATTCTTCCTTCATGCTGAAACAACAATCC 412
 QY 21 ArgAaGleuGlnHisIsthPheValPheLeuAaGaaPheSerLeuMetLeuLeuAaGly 40
 Db 413 AGAAGGCTGGAAACAATACCTTGTCTCTTGAAGAAATTTTCCAGATATTATTAAAGATAC 472

Oy		41	IIIGLYVSLYVSARGAAGALATHNARvPhETripAPROARGAGLThPro	58
Dd		473	ATTGGCAAGAAAGAAGACACGATTTCTGGATCCAGAGGGACAACCA	526
RESULT 39				
AX332835				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
ORIGIN				
Alignment Scores:				
Pred. No.:				
Score:				
Percent Similarity:				
Best Local Similarity:				
Query Match:				
DB:				
US-10-031-158-14 (1-58)	x	AX332835 (1-1421)		
Oy		1	MelGlnMetPhePrOFProSerFProLeuPhePheLeuGlnLeuLeuGlnSerSer	20
Dd		423	ATGAGATGTTCCTCCCCAACGCCACTATTTTCTTCCTCAATTCGAAACAAAACCTCC	482
Oy		21	ArgArgLeuGluIuhstHrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTYR	40
Dd		483	AGAAAGCTCGGAACATCACTTGTTGCTTCCTTGGAATTTTTCCACAGATATTTAAGATAC	542
Oy		41	IIIGLYVSLYVSARGAAGALATHNARvPhETripAPROARGAGLThPro	58
Dd		543	ATTGGCAAGAAAGAAGACACGATTTCTGGATCCAGAGGGACAACCA	596
RESULT 40				
HUMTCGXH				
LOCUS				
DEFINITION				
ACCESION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				

COMMENT

Original source text: Human T-cell line HPB-MLT, cDNA to mRNA,
 clones pT-gamma-[1,2].
 Computer-readable sequence of [1] kindly provided by D.P. Dialynas,
 31-Oct-1986.
 An aberrant rearrangement at the V-J junction causes a frame shift
 that leads to peptide termination at the 'taa' codon at positions
 370-372.

FEATURES

A polyadenylation signal can be found at positions 1400 to 1405.
 Location/Qualifiers

source

1..1421
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

CDS

4..979
 /note="pseudo-TCG pept"
 /pseudo
 /codon_start=1
 4..63
 /note="pseudo-TCG signal pept"
 /pseudo
 64..976
 /note="pseudo-TCG mature pept"
 /pseudo

sig_peptide

mat_peptide

ORIGIN 16 bp upstream of ApaI site; chromosome 7.

Alignment Scores:

Pred. No.:	6.72e-28	Length:	1421
Score:	294.00	Matches:	55
Percent Similarity:	98.28%	Conservative:	2
Best Local Similarity:	94.83%	Mismatches:	1
Query Match:	95.77%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x HUMTCGNH (1-1421)

QY	1	MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer	20
DB	423	ATGCAGATGTTCCCCCAAGCCCACTATTCTTCCTTCAATGCTGAAACAAACTCC	482
QY	21	ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr	40
DB	483	AGAAAGCTGGAACATACCTTGTGCTTCTTGAGAAATTTTCCAGATATTATTAAAGATAC	542
QY	41	IleGlyValValArgArgAlaThrArgPheTTPaapProArgArgGlyThrPro	58
DB	543	ATTGGCAAGAAAGAGAGACACGATCTGGGATCCCGAGAGGAGGAAACACCA	596

Search completed: December 8, 2004, 08:54:17
 Job time : 1916 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2004, 06:09:26 ; Search time 247 Seconds
(without alignments)
1232.658 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307

Sequence: 1 MQMFPSPDLFFFLQLKQSS.....RYGKKRATRFMPRRGTP 58

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_epool/US10031158/runat_06122004_082658_15877/app_query.fasta_1.199
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=40 -ALIGN=40
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158 -CGCN_1_1_352 @runat_06122004_082658_15877 -NCPU=6 -ICPU=3
-NO KWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_23Sep04:*
1: geneseqn19808:*
2: geneseqn19808:*
3: geneseqn20008:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	426	3 AAC00681	AAC00681 Human sec
2	307	100.0	470	4 AAK26982	AAK26982 Human bon
3	307	100.0	470	6 ABS01569	ABS01569 Human gen
4	307	100.0	477	4 ABA43588	ABA43588 Human bre
5	307	100.0	477	4 AAK27758	AAK27758 Human bon
6	307	100.0	477	6 ABS02210	ABS02210 Human gen

c	7	307	100.0	539	5	ABV28618	ABV28618 Human pro
c	8	307	100.0	539	2	ABV22791	ABV22791 Human pro
c	9	307	100.0	825	5	AAO66888	AAO66888 Human lym
c	10	307	100.0	825	9	ACF35988	ACF35988 Human T c
c	11	307	100.0	1027	4	AAFS6401	AAFS6401 Human TCR
c	12	307	100.0	1027	10	ADB75580	ADB75580 Prostate
c	13	307	100.0	1316	5	ABV26485	ABV26485 Human pro
c	14	307	100.0	1316	5	ABV25093	ABV25093 Human pro
c	15	307	100.0	1395	12	ADL06443	ADL06443 Human tum
c	16	307	100.0	1418	6	AAD38828	AAD38828 Human PSN
c	17	307	100.0	2658	10	ADE08798	ADE08798 Novel DNA
c	18	304	99.0	486	5	ABV25013	ABV25013 Human pro
c	19	302	98.4	316	4	ABA47985	ABA47985 Human bre
c	20	302	98.4	316	4	AAK40015	AAK40015 Human bon
c	21	302	98.4	316	6	ABSI4101	ABSI4101 Human gen
c	22	298	97.1	569	5	ABV43475	ABV43475 Human pro
c	23	294	95.8	1421	6	ABL65007	ABL65007 Lung cancr
c	24	294	95.8	1560	10	ADE06987	ADE06987 Novel cod
c	25	294	95.8	1799	6	ABZ35381	ABZ35381 Human gen
c	26	290	94.5	1080	2	AAQ37617	AAQ37617 TCR gamma
c	27	290	94.5	2658	10	ADE08798	ADE08798 Novel DNA
c	28	284	92.5	1046	12	ADP10449	ADP10449 Reference
c	29	284	92.5	1586	6	ABK64529	ABK64529 Human den
c	30	279	90.9	1080	5	AAV91698	AAV91698 Sequence
c	31	259	84.4	533	5	ABV37856	ABV37856 Human pro
c	32	251	81.8	460	3	AAQ04122	AAQ04122 Human sec
c	33	245	79.8	1155	3	AAFI5623	AAFI5623 Human pro
c	34	240	78.2	1022	5	ABV29989	ABV29989 Human pro
c	35	240	78.2	1022	5	ABV24072	ABV24072 Human pro
c	36	240	78.2	1022	5	ABV24124	ABV24124 Human pro
c	37	240	78.2	1022	5	ABV25823	ABV25823 Human pro
c	38	240	78.2	1022	5	ABV24034	ABV24034 Human pro
c	39	240	78.2	1022	5	ABV23986	ABV23986 Human pro
c	40	240	78.2	1022	5	ABV23991	ABV23991 Human pro
c	41	240	78.2	1022	5	ABV24108	ABV24108 Human pro
c	42	240	78.2	1022	5	ABV23814	ABV23814 Human pro
c	43	240	78.2	1022	5	ABV29918	ABV29918 Human pro
c	44	240	78.2	1022	5	ABV29874	ABV29874 Human pro
c	45	240	78.2	1022	5	ABV29953	ABV29953 Human pro

ALIGNMENTS

RESULT 1	
AA000681	standard; cDNA; 426 BP.
ID	AA000681 standard; cDNA; 426 BP.
AC	AA000681;
XX	
AC	AA000681;
XX	
DT	06-OCT-2000 (first entry)
XX	
DE	Human secreted protein 5' EST, SEQ ID NO: 679.
XX	
DE	Human; 5' EST, expressed sequence tag; secreted protein; cDNA isolation;
XX	
KW	gene therapy; chromosome mapping; ss.
XX	
OS	Homo sapiens.
XX	
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-00200610.
XX	
PR	26-FEB-1999; 99US-0122487P.
XX	
PA	(GENSET) GENSET.
XX	
PI	Dumas Milne Edwards J, Duclert A, Giordano J;
XX	
DR	WPI; 2000-500381/45.
XX	
DR	P-PSDB; AAG00675.
XX	

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1, SEQ ID NO 679; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors

XX SQ Sequence 426 BP; 151 A; 83 C; 88 G; 104 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4, 87e-34	426	58	0	0	0	0
Score:	307.00	58	0	0	0	0
Percent Similarity:	100.00%	100.00%	0	0	0	0
Best Local Similarity:	100.00%	100.00%	0	0	0	0
Query Match:	3	3	0	0	0	0

US-10-031-158-14 (1-58) x AAC00681 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20

DB 94 ATGCAAGATGTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAAACAAAGCTCC 153

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40

DB 154 AGAAGGCTGGACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 213

QY 41 TleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 214 ATTGGCAAGAAAGAAAGACACGATCTGGGATCCCGAGGAGGAACACCA 267

RESULT 2

AKK26982 ID AKK26982 standard; DNA; 470 BP.

XX AAK26982;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 1539.

KM Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

XX WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

PS Example 4; SEQ ID NO 1539; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention

XX SQ Sequence 470 BP; 148 A; 102 C; 87 G; 133 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5, 56e-34	470	58	0	0	0	0
Score:	307.00	58	0	0	0	0
Percent Similarity:	100.00%	100.00%	0	0	0	0
Best Local Similarity:	100.00%	100.00%	0	0	0	0
Query Match:	4	4	0	0	0	0

US-10-031-158-14 (1-58) x AAK26982 (1-470)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSer 20

DB 203 ATGCAAGATGTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAAACAAAGCTCC 262

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40

DB 263 AGAAGGCTGGACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 322

QY 41 TleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 323 ATTGGCAAGAAAGAAAGACACGATCTGGGATCCCGAGGAGGAACACCA 376

RESULT 3

ABS01569 ID ABS01569 standard; DNA; 470 BP.

XX ABS01569;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID NO 1560.

KM Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KM chronic obstructive pulmonary disease; interstitial lung disease;

KM familial idiopathic pulmonary fibrosis; neurofibromatosis;

KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KM Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;

KM pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;

KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KM primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX Claim 1; SEQ ID NO 1560; 634pp; English.
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridize at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarray having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Judd syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 470 BP; 148 A; 102 C; 87 G; 133 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,56e-34 Length: 470
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-031-158-14 (1-58) x ABS01569 (1-470)
 QY 1 MetGImetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
 DB 203 ATGCAATGATGTTCCCAAGCCACATATTTCTTCCTCAATGCTGAACAAAGCTCC 262
 QY 21 AAGATGLeuGlnHisThrPheValPheLeuArgAsnPhSerSerLeuLeuArgTyr 40

DB 263 AGAAGGCTGAACATACCTTGTCTTGTGAGAAATTTTCCGTATGTTAAGATAC 322
 QY 41 TleGlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58
 DB 323 ATTGGCAAGAAAGAAAGAGCAACAGATTCTGGATCCAGAGGGGAACACCA 376
 RESULT 4
 ID ABA43588 standard; DNA; 477 BP.
 XX ABA43588;
 AC ABA43588;
 XX 01-FEB-2002 (first entry)
 DT
 XX Human breast cell single exon nucleic acid probe #2283.
 DE Human; microarray; single exon probe; gene expression; breast; disease;
 KM cancer; ss.
 XX Homo sapiens.
 OS WO200157271-A2.
 PN 09-AUG-2001.
 PD 30-JUN-2001; 2001WO-US000662.
 PF 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207455P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-496933/54.
 PT New spatially-addressable set of single exon nucleic acid probes, useful
 PT for measuring gene expression in sample derived from human breast, useful
 PT comprises number of single exon nucleic acid probes.
 XX Claim 1; SEQ ID NO 2283; 327pp + Sequence Listing; English.
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC prognosing breast disease. Gene expression analysis is useful for
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a single exon
 CC nucleic acid probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 477 BP; 147 A; 104 C; 88 G; 138 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,67e-34 Length: 477
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x ABA43588 (1-477)

Oy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20
 |||||
 Db 212 ATGCAGATGTTTCCCCCAAGCCACATATTTCTTCTCAATGCTGAACAAAGCTCC 271
 |||||

Oy 21 ArgArgLeuGluHisIleThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
 |||||
 Db 272 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCGTGATGTATTATAAGATAC 331
 |||||

Oy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 |||||
 Db 332 ATTGGCAAGAAAGAAAGAGACACATCTGGGATCCACAGAGGGAACACCA 385
 |||||

RESULT 5
 AAK27758
 ID AAK27758 standard; DNA; 477 BP.
 XX
 AC AAK27758;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed single exon probe SEQ ID NO: 2315.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukemia; lymphoma; myeloma; ss.
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000668.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX
 PS gene expression in human bone marrow.
 XX
 PS Example 4; SEQ ID NO 2315; 658bp + Sequence listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention
 XX
 SQ Sequence 477 BP; 147 A; 104 C; 88 G; 138 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5.67e-34 Length: 477
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0
 US-10-031-158-14 (1-58) x AAK27758 (1-477)

Oy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20
 |||||
 Db 212 ATGCAGATGTTTCCCCCAAGCCACATATTTCTTCTCAATGCTGAACAAAGCTCC 271
 |||||

Oy 21 ArgArgLeuGluHisIleThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
 |||||
 Db 272 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCGTGATGTATTATAAGATAC 331
 |||||

Oy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 |||||
 Db 332 ATTGGCAAGAAAGAAAGAGACACATCTGGGATCCACAGAGGGAACACCA 385
 |||||

RESULT 6
 ABS02210
 ID ABS02210 standard; DNA; 477 BP.
 XX
 AC ABS02210;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe from lung SEQ ID NO 2201.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KM Hermansky-Rudrak syndrome; sarcoidosis; pulmonary haemorrhosis;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 XX
 PS measure gene expression in human lung samples.
 XX
 PS Claim 1; SEQ ID NO 2201; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC and the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 1201 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Heremakky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 477 BP; 147 A; 104 C; 88 G; 138 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,67e-34 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x ABS02210 (1-477)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20
DB 212 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCAATGCTGAAACAAGCTCC 271
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 272 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 331
QY 41 IleglyLyLyArgArgAlaThrArgPheTTPAAPPProArgArgGlyThrPro 58
DB 332 ATTGGCAAGAAAGAAAGAGCAACGATTCCTGGATCCAGAGGGGAACACCA 385

RESULT 7

ABV28618/C
ID ABV28618 standard; cDNA, 539 BP.

XX AC ABV28618;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 28609.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183131P.
PR 16-MAR-2000; 2000US-0189863P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endage WO, Monahan JE;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5996; 11750p; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 539 BP; 134 A; 104 C; 106 G; 195 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,68e-34 Length: 539
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x ABV28618 (1-539)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20
DB 410 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCAATGCTGAAACAAGCTCC 351
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 350 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 291
QY 41 IleglyLyLyArgArgAlaThrArgPheTTPAAPPProArgArgGlyThrPro 58
DB 290 ATTGGCAAGAAAGAAAGAGCAACGATTCCTGGATCCAGAGGGGAACACCA 237

RESULT 8

ABV22791/C
ID ABV22791 standard; cDNA, 539 BP.

XX AC ABV22791;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 22782.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.

XX 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183119P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0235281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Bodege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 4022; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (1) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement; (1) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC
 XX
 SQ Sequence 539 BP; 134 A; 104 C; 106 G; 195 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6,68e-34 Length: 539
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-031-158-14 (1-58) x ABV22791 (1-539)
 QY 1 MetG1MePhePrOPrOsErPrOleuPhePhePheLeuG1nLeuLeuYsGInSerSer 20
 DB 410 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTCGAAACAAAGCTCC 351
 QY 21 ArgArgLeuG1uH1sThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTYR 40
 DB 350 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCGATGTTATTAAAGTAC 291
 QY 41 IleG1LyLeuYsArgArgAlaThrArgPheTrpAspProArgArgG1YThrPro 58
 DB 290 ATTGGCAAGAAAGAAAGACACACGATTCGGATCCACAGAGGGAACACCA 237
 RESULT 9
 ID AAQ66888 standard; cDNA; 825 BP.
 XX
 AC AAQ66888;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-FEB-1995 (first entry)
 XX
 DE Human lymphocyte clone G 115 soluble TCR Vgamma chain.
 XX
 KM Soluble T cell receptor; TCR; pathological T lymphocyte; PCR primer;

KW polymerase chain reaction; leukaemia; Burkitt's lymphoma;
 KW autoimmune disease; cell typing; gamma; delta; lymphocyte g115;
 KW Mycobacterium tuberculosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..825
 FT /*tag= a
 FT /product= "soluble_Vgamma_chain"
 XX
 EN WO9412648-A2.
 XX
 PD 09-JUN-1994.
 XX
 PF 25-NOV-1993; 93WO-FR001165.
 XX
 PR 25-NOV-1992; 92FR-00014203.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (IMMU-) IMMUNOTECH SA.
 XX
 PI Bonneville M;
 XX
 DR WPI; 1994-200276/24.
 DR P-PSDB; AAR5705.
 XX
 PT Prodn. of soluble T receptors lacking the trans-membrane region - by co-
 PT transfection of host cells with truncated sub unit DNA, for diagnosis and
 PT treatment of pathological T cell proliferation.
 XX
 PS Example 1; Fig 2B; 41pp; French.
 XX
 CC The human lymphocyte clone gamma-delta-s G115 expresses T cell receptors
 CC V9Jr3gamma/V2D3J1delta. The soluble delta and gamma chains have the
 CC sequences AAQ66887 and AAQ66888, respectively. This clone, which kills
 CC Burkitt's lymphoma ("Dad1") cells and recognises an antigen present in
 CC water-soluble extracts of M.tuberculosis, was used for the construction
 CC of gamma-delta-s genes and expression of soluble T cell receptors. The
 CC soluble TCRs (i.e. lacking the transmembrane region) are useful in
 CC diagnosis (esp. cell typing) and treatment of conditions involving T cell
 CC proliferation, e.g. leukaemia, lymphoma or some autoimmune diseases.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 825 BP; 267 A; 178 C; 176 G; 204 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.19e-33 Length: 825
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-031-158-14 (1-58) x AAQ66888 (1-825)
 QY 1 MetG1MePhePrOPrOsErPrOleuPhePhePheLeuG1nLeuLeuYsGInSerSer 20
 DB 440 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTCGAAACAAAGCTCC 499
 QY 21 ArgArgLeuG1uH1sThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTYR 40
 DB 500 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCGATGTTATTAAAGTAC 559
 QY 41 IleG1LyLeuYsArgArgAlaThrArgPheTrpAspProArgArgG1YThrPro 58
 DB 560 ATTGGCAAGAAAGAAAGACACACGATTCGGATCCACAGAGGGAACACCA 613
 RESULT 10
 ID ACF35988 standard; cDNA; 825 BP.
 XX
 AC ACF35988;

[illegible]

Db	500	AGAAAGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCTGATGTTATTAAAGATAC	595
Qy	41	IIegIyLbYbLbSArGArGAlaThArGPhetTPaAPProArGlyThrPro	58
Db	560	ATTGGAGAGAAAGAAAGAGCAACAGATTCTGGGATCCGAGAGGGGAACACCA	613
RESULT 11			
ID	AAFS6401	standard; cDNA, 1027 BP.	
AC	AAFS6401;		
XX	12-APR-2001	(first entry)	
DE	Human TCRgamma alternate reading frame protein coding sequence.		
KW	Human; TARP; prostate cancer; breast cancer; immunotherapy;		
XX	T cell receptor gamma alternate reading frame protein; TCRgamma; ss.		
OS	Homo sapiens.		
PN	WO200104309-A1.		
PD	18-JAN-2001.		
PF	12-JUL-2000; 2000WO-US019039.		
PR	13-JUL-1999; 99US-0143560P.		
PR	01-OCT-1999; 99US-0157471P.		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PA	Paetan I, Essand M, Lee B, Vasmatazis G, Wolfgang C;		
PI	MP1; 2001-081050/09.		
DR	P-P8DB; AAB66399, AAB66400.		
XX	Isolated T-cell receptor gamma alternate reading frame protein useful for		
PT	diagnosing and raising an immune response to prostate cancer and breast		
XX	cancer.		
XX	Example 1; Fig 1; 85dp; English.		
CC	The present invention provides the protein and coding sequences of the		
CC	human T cell receptor alternate reading frame protein (TARP). This		
CC	protein is expressed in prostate and breast cancer cells at higher levels		
CC	than normal and so can be used in the immunotherapy of these cancers, as		
XX	well as their detection and prevention		
XX	Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;		
Alignment Scores:			
Prod. No.:	1-59e-33	Length:	1027
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0
US-10-031-158-14 (1-58) x AAF56401 (1-1027)			
Qy	1	MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuYsgInserSer	20
Db	74	ATGCAGATGTTTCCCCCAAGCCCATATTTTCTTCCTCAATGCTGAACAAAGCTCC	133
Qy	21	ArgAagLeuGlnHisThrPheValPheLeuAgaenPheSerLeuMetLeuLeuAgtTyr	40
Db	134	AGAAAGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC	193
Qy	41	IIegIyLbYbLbSArGArGAlaThArGPhetTPaAPProArGlyThrPro	58
Db	194	ATTGGAGAGAAAGAAAGAGCAACAGATTCTGGGATCCGAGAGGGGAACACCA	247

RESULT 12
 ADB75580
 ID ADB75580 standard; cDNA; 1027 BP.
 XX
 AC ADB75580;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Prostate cancer marker cDNA.
 XX
 KM Prostate; cancer; cytostatic; gene therapy; marker; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003009814-A2.
 XX
 PD 06-FEB-2003.
 XX
 PF 25-JUL-2002; 2002WO-US023913.
 XX
 PR 25-JUL-2001; 2001US-0307982P.
 PR 22-AUG-2001; 2001US-0314356P.
 PR 25-SEP-2001; 2001US-0325020P.
 PR 12-DEC-2001; 2001US-0341746P.
 PR 05-MAR-2002; 2002US-0362158P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbatcheva B;
 PI Hoersch S, Kamatkar S, Womsey AM, Glatz K, Zhao X, Anderson D;
 DR WPI; 2003-248033/24.
 XX
 PT New nucleic acid molecule, useful for diagnosing or treating prostate
 PT cancer.
 XX
 PS Disclosure; SEQ ID NO 404; 99pp; English.
 XX
 CC The invention relates to newly discovered cancer markers associated with
 CC the cancerous state of prostate cells. Also disclosed is a method of
 CC assessing whether a patient is afflicted with prostate cancer. The method
 CC of the invention involves assessing whether a patient is afflicted with
 CC prostate cancer by comparing the level of expression of a marker in a
 CC patient sample and the normal level of expression of the marker in a
 CC control non-prostate cancer sample, where a significant increase in the
 CC level of expression of the marker in the patient sample and the normal
 CC level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating
 CC prostate cancer, and may be useful in gene therapy. Sequences given in
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: 1,59e-33 Length: 1027
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-031-158-14 (1-58) x ADB75580 (1-1027)

QY 1 MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20
 DB 74 ATTCAGATGTTTCCCAAGCCACATATTTTCTTCTTCAATTCGTAAGAAAGCTCC 133

QY 21 ArgArgLeuGlnLuhIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
 DB 134 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 193

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
 DB 194 ATTGGCAAGAAAAGAAAGAGCAACGATTTCTGGATCCACGAGGAGCAACACCA 247

RESULT 13
 ABV26485/c
 ID ABV26485 standard; cDNA; 1316 BP.
 XX
 AC ABV26485;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 26476.
 XX
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KM pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0215007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 PI WPI; 2001-662795/76.
 XX
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 5351; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX

Alignment Scores:
 Pred. No.: 2,22e-33 Length: 1316
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x ABV26485 (1-1316)

QY 1 MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20
 DB 562 ATTCAGATGTTTCCCAAGCCACATATTTTCTTCTTCAATTCGTAAGAAAGCTCC 503

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAspHisSerLeuMetLeuLeuArgTyr 40
 Db 502 AGAAGGCTGGAACATACCTTTCTCTTGAAGAAATTTTCCCTGATGTATTAAGATAC 443

Qy 41 IleGlyIysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 442 ATTGGCAAGAAAGAAAGAGACACGATTTCTGGATCCAGAGGGGAAACACCA 389

RESULT 14
 ABV25093/c
 ID ABV25093 standard; cDNA; 1316 BP.
 AC ABV25093;
 XX
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker CDNA 25084.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 XX WO200160860-A2.
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-021907P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 PS Claim 1; Page 4881; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 XX SQ Sequence 1316 BP; 369 A; 282 C; 300 G; 365 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2, 22e-33 Length: 1316
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0
 US-10-031-158-14 (1-58) x ABV25093 (1-1316)

Qy 1 MetGlnMetPheProProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
 Db 562 ATGCAGATGTTTCCGCCAAGCCCACTAATTTTCTTCCCTGATGTGTAACAAAGCTCC 503

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAspHisSerLeuMetLeuLeuArgTyr 40
 Db 502 AGAAGGCTGGAACATACCTTTCTCTTGAAGAAATTTTCCCTGATGTATTAAGATAC 443

Qy 41 IleGlyIysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 442 ATTGGCAAGAAAGAAAGAGACACGATTTCTGGATCCAGAGGGGAAACACCA 389

RESULT 15
 ADL06443
 ID ADL06443 standard; cDNA; 1395 BP.
 AC ADL06443;
 XX
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #23.
 XX
 KW Human; tumour-associated antigenic target; TAT; cell death; tumour;
 KW cancer; cytostatic; gene; ss.
 OS Homo sapiens.
 XX
 XX WO2004016225-A2.
 PN 26-FEB-2004.
 PD 19-AUG-2003; 2003WO-US025892.
 XX
 PF 19-AUG-2002; 2002US-0404809P.
 PR 21-AUG-2002; 2002US-0405645P.
 PR 23-SEP-2002; 2002US-0413192P.
 PR 15-OCT-2002; 2002US-0419008P.
 PR 15-NOV-2002; 2002US-0426847P.
 PR 02-JUL-2003; 2003US-0484959P.
 XX
 PA (GETH) GENENTECH INC.
 PI Deesavage RJ, Franz G, Hillan KJ, Polakis P, Poleon A, Smith V;
 PI Spencer SD, Wu TD, Zhang Z;
 DR WPI; 2004-257144/24.
 DR P-PsDB; ADL06523.

CC New antibody that binds to a tumor-associated antigenic target (TAT)
 CC polypeptide, useful for preparing a composition for diagnosing or
 CC treating cancer.
 PS Claim 1; SEQ ID NO 23; 319pp; English.

CC The present invention relates to the isolation of human tumour-associated
 CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also
 CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
 CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
 CC humanized antibody. It is conjugated to a growth inhibitory agent. It is
 CC produced in bacteria or in CHO cells and induces death of a cell to which
 CC it binds. The antibody is useful for preparing a composition for
 CC diagnosing or treating tumours and cancer. The present sequence
 CC represents a human TAT cDNA sequence of the invention.
 XX
 XX SQ Sequence 1395 BP; 405 A; 324 C; 286 G; 380 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2, 41e-33 Length: 1395
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-10-031-158-14 (1-58) x ADL06443 (1-1395)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInSer 20
 |||
 DB 449 ATGCAGATGTTTCCCAAGCCACGATTTTCTTCAATGCTGAACAAAGCTCC 508

QY 21 ATGATGLeuGlnHleThPheValPheLeuArgAsnPheSerLeuMetLeuArgTyr 40
 |||
 DB 509 AGAAGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCGTATGTTATTAAAGATAC 568

QY 41 TleGlyLysLysArgArgAlaThArgPheTTPaPProArgArgGlyThrPro 58
 |||
 DB 569 ATTGGCAAGAAAAGAGACCAACGATTCGGGATCCAGAGGGGAACACCA 622

RESULT 16

AD38828
 ID AAD38828 standard; cDNA; 1418 BP.

AC AAD38828;
 XX
 XX 23-SEP-2002 (first entry)
 DE Human PSNA CDNA #8.
 XX
 XX Human; prostate specific protein; PSP; prostate specific nucleic acid;
 KM vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;
 KM PSNA; 88.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200224718-A1.
 PN
 XX 28-MAR-2002.
 PD
 XX 19-SEP-2001; 2001WO-US023386.
 PF
 XX 19-SEP-2000; 2000US-0233746P.
 PR
 XX (DIAD-) DIADEXUS INC.
 PA
 XX Sun Y, Recipon H, Cafferty R, Ali S;
 PI
 XX WPI; 2002-471216/50.
 DR
 XX
 XX Novel isolated prostate specific polypeptide useful for identifying,
 PT diagnosing, monitoring, staging, imaging, and treating prostate cancer
 PT and non-cancerous disease states in prostate.
 PT
 XX
 XX Claim 1; Page 187; 210pp; English.
 PS
 XX The invention relates to prostate specific proteins (PSP) and prostate
 CC specific nucleic acids (PSNA). Sequences of the invention are useful for
 CC identifying, diagnosing, monitoring, staging, imaging and treating
 CC prostate cancer and non-cancerous disease states in prostate. They are
 CC also useful for producing engineered prostate tissue for treatment and
 CC research. The PSNA sequences are used in gene therapy and for producing
 CC transgenic animals and cells. The invention is also used as vaccines. The
 CC present sequence is human PSNA CDNA
 CC
 XX
 SQ Sequence 1418 BP; 418 A; 335 C; 285 G; 380 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,46e-33 Length: 1418
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AAD38828 (1-1418)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInSer 20
 |||
 DB 449 ATGCAGATGTTTCCCAAGCCACTATTTTCTTCAATGCTGAACAAAGCTCC 508

QY 21 ATGATGLeuGlnHleThPheValPheLeuArgAsnPheSerLeuMetLeuArgTyr 40
 |||
 DB 509 AGAAGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCGTATGTTATTAAAGATAC 568

QY 41 TleGlyLysLysArgArgAlaThArgPheTTPaPProArgArgGlyThrPro 58
 |||
 DB 569 ATTGGCAAGAAAAGAGACCAACGATTCGGGATCCAGAGGGGAACACCA 622

RESULT 17

ADE08798/C
 ID ADE08798 standard; DNA; 2658 BP.

AC ADE08798;
 XX
 XX 29-JAN-2004 (first entry)
 DE Novel DNA-related contig nucleotide sequence #42.
 XX
 XX novel gene; novel protein; tissue marker; molecular weight marker;
 KM chromosome marker; genetic disorder; contig; ds.
 KM
 XX
 XX Unidentified.
 OS
 XX
 XX WO2003054152-A2.
 PN
 XX 03-JUL-2003.
 PD
 XX 10-DEC-2002; 2002WO-US039555.
 PF
 XX
 XX 10-DEC-2001; 2001US-0339739P.
 PR
 XX 11-DEC-2001; 2001US-0339453P.
 PR
 XX 14-MAR-2002; 2002US-0365091P.
 PR
 XX 14-MAR-2002; 2002US-0365384P.
 PR
 XX 12-APR-2002; 2002US-0372381P.
 PR
 XX 12-APR-2002; 2002US-0372615P.
 PR
 XX 22-APR-2002; 2002US-00128558.
 PR
 XX 24-APR-2002; 2002US-0376045P.
 PA
 XX (HYSE-) HYSEQ INC.
 PI
 XX Tang YT, Asundi V, Goodrich RM, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Meng G, Zhou P, Drmanac RT, Wang Z;
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ,
 DR WPI; 2003-569235/53.
 XX
 XX New polynucleotides, useful for expressing recombinant proteins for
 PT analysis, characterization or therapeutic use, or as markers for tissues
 PT in which the corresponding protein is preferentially expressed.
 PT
 XX
 XX Disclosure; SEQ ID NO 1864; 1177pp; English.
 PS
 XX The invention comprises the amino acid and coding sequences of novel
 CC proteins. The DNA and protein sequences of the invention are useful as:
 CC markers for tissues in which the corresponding protein is preferentially
 CC expressed; as molecular weight markers on gels; as chromosome markers or
 CC tags; to identify chromosomes or to map related gene positions; and to
 CC compare with endogenous DNA sequences in patients to identify potential
 CC genetic disorders. The present DNA sequence was used in the
 CC exemplification of the invention.
 CC
 XX
 SQ Sequence 2658 BP; 745 A; 585 C; 578 G; 750 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,74e-33 Length: 2658
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-10-031-158-14 (1-58) x ADE08798 (1-2658)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuIleuGlnSerSer 20

DB 2381 ATGCAGATGTTCCCGCAAGCCACTATTTCTTCAATTCCTGAAACAAAGCTCC 2322

QY 21 ArgArgLeuGlnIleThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 2331 AGAAGGCTGGAACACACTTTGTCTTCTTGAGAAAATTTTCCCTGATGTTATTAAAGATAC 2262

QY 41 IleGlyIleValArgArgAlaThrArgPheTTPAAPPProArgGlyThrPro 58

DB 2261 ATTGCAAGAAAGAAAGACACACGATTCGCGATCCAGAGGGGAACACCA 2208

RESULT 18

ABV25013/C

ID ABV25013 standard; cDNA; 466 BP.

XX AC ABV25013;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker CDNA 25004.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KM pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001:

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 4846; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (1) comprising

XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX CC specification or its complement. (1) is useful for: (a) assessing whether

XX CC a patient is afflicted with prostate cancer; (b) monitoring the

XX CC progression of prostate cancer in a patient; (c) assessing the efficacy

XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX CC determining whether prostate cancer has metastasized in a patient; (h)

XX CC assessing the aggressiveness or indolence of prostate cancer in a patient

XX CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 466 BP; 117 A; 94 C; 101 G; 174 T; 0 U; 0 Other;

Alignment Scores: 1.55e-33 Length: 466

Pred. No.: 1.55e-33 Length: 466

Score: 304.00 Matches: 57

Percent Similarity: 100.00% Conservatve: 1

Best Local Similarity: 98.28% Mismatches: 0

Query Match: 99.02% Indels: 0

DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x ABV25013 (1-486)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuIleuGlnSerSer 20

DB 411 ATGCAGATGTTCCCGCAAGCCACTATTTCTTCAATTCCTGAAACAAAGCTCC 352

QY 21 ArgArgLeuGlnIleThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 351 AGAAGGCTGGAACACACTTTGTCTTCTTGAGAAAATTTTCCCTGATGTTATTAAAGATAC 292

QY 41 IleGlyIleValArgArgAlaThrArgPheTTPAAPPProArgGlyThrPro 58

DB 291 ATTGCAAGAAAGAAAGACACGATTCGCGATCCAGAGGGGAACACCA 238

RESULT 19

ABA47985

ID ABA47985 standard; DNA; 316 BP.

XX AC ABA47985;

XX DT 01-FEB-2002 (first entry)

XX DE Human breast cell single exon nucleic acid probe #6680.

XX KW Human; microarray; single exon probe; gene expression; breast; disease;

XX KM cancer; ss.

XX OS Homo sapiens.

XX PN WO200157271-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000662.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-496933/54.

XX PT New spatially-addressable set of single exon nucleic acid probes, useful

XX PT for measuring gene expression in sample derived from human breast,

XX PT comprises number of single exon nucleic acid probes.

XX PS Claim 4; SEQ ID NO 6680; 327pp + Sequence Listing; English.

XX CC The invention relates to a spatially-addressable set of single exon

XX CC nucleic acid probes for measuring gene expression in a sample derived

XX CC from human breast and BT 474 cells. The method involves contacting the

XX CC probes with a collection of detectably labelled nucleic acids derived

XX CC from mRNA of human breast, and then measuring the label bound to each

XX CC probe of the microarray. The probes are useful for verifying the

XX CC expression of regions of genomic DNA predicted to encode proteins. They

XX CC are useful for gene discovery, and for determining predisposition and/or

XX CC prognosing breast disease. Gene expression analysis is useful for

XX CC assessing the toxicity of chemical agents on cells. The microarray of

XX CC this invention presents a far greater diversity of probes for measuring

XX CC gene expression, with far less bias than expressed sequence tag

CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcc_sequences

XX SQ Sequence 316 BP; 112 A; 65 C; 62 G; 77 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,67e-33	Length:	316
Score:	302.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.37%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x ABA47985 (1-316)

QY 2 GlnMetPheProPseSerProLeuPhePheLeuGlnLeuLySGlnSerSerArg 21
|||
DB 2 CAGATGTTTCCCCCAAGCCCACTATTCTTCTTCAATTGCTGAACAAAGCTCCACA 61
|||
QY 22 ArgLeuGlnHisThrPheValPheLeuArgaenPheSerLeuMetLeuArgTyrIle 41
|||
DB 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTAATTAAGATACATT 121
|||
QY 42 GtLyLySArGArGAlaThrArPheTTPaEPProArGArGtYThrPro 58
|||
DB 122 GGCAGAAAGAAAGAGACACGATTCTGGATCCAGAGAGGAAACACCA 172
|||

RESULT 20

AAK40015
ID AAK40015 standard; DNA; 316 BP.

XX AAK40015;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed single exon probe SEQ ID NO: 14572.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX OS Homo sapiens.

XX PN MO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001MO-US000668.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX Example 4; SEQ ID NO 14572; 658bp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention

XX SQ Sequence 316 BP; 112 A; 65 C; 62 G; 77 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,67e-33	Length:	316
Score:	302.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.37%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x AAK40015 (1-316)

QY 2 GlnMetPheProPseSerProLeuPhePheLeuGlnLeuLySGlnSerSerArg 21
|||
DB 2 CAGATGTTTCCCCCAAGCCCACTATTCTTCTTCAATTGCTGAACAAAGCTCCACA 61
|||
QY 22 ArgLeuGlnHisThrPheValPheLeuArgaenPheSerLeuMetLeuArgTyrIle 41
|||
DB 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTAATTAAGATACATT 121
|||
QY 42 GtLyLySArGArGAlaThrArPheTTPaEPProArGArGtYThrPro 58
|||
DB 122 GGCAGAAAGAAAGAGACACGATTCTGGATCCAGAGAGGAAACACCA 172
|||

RESULT 21

ABS14101
ID ABS14101 standard; DNA; 316 BP.

XX ABS14101;

XX DT 19-AUG-2002 (first entry)

XX DE Human genome-derived single exon probe ORF from lung SEQ ID NO 14092.

XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX KM familial idiopathic pulmonary fibrosis; interstitial lung disease;

XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX KM Hereditary Pulchak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;

XX KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX KW primary ciliary dyskinesia; pulmonary hypertension;

XX KM hyaline membrane disease; open reading frame; ORF.

XX OS Homo sapiens.

XX PN MO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001MO-US000665.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX

XX	AC	ABL65007;	
XX	DT	15-MAY-2002	(first entry)
XX	XX		
DE	XX		Lung cancer related gene sequence SEQ ID NO:3344.
KM	XX		Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KM	XX		stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW	XX		cytostetic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX	OS		gene; ds.
XX	XX		
XX	XX		Homo sapiens.
XX	XX		
XX	PN		WO200194629-A2.
XX	PD		13-DEC-2001.
XX	XX		
XX	PF		30-MAY-2001; 2001WO-US010838.
XX	XX		
PR	XX		05-JUN-2000; 2000US-0209473P.
PR	XX		05-JUN-2000; 2000US-0209531P.
PR	XX		18-SEP-2000; 2000US-0233133P.
PR	XX		18-SEP-2000; 2000US-0233617P.
PR	XX		20-SEP-2000; 2000US-0234009P.
PR	XX		20-SEP-2000; 2000US-0234034P.
PR	XX		20-SEP-2000; 2000US-0234052P.
PR	XX		22-SEP-2000; 2000US-0234509P.
PR	XX		22-SEP-2000; 2000US-0234567P.
PR	XX		25-SEP-2000; 2000US-0234923P.
PR	XX		25-SEP-2000; 2000US-0234924P.
PR	XX		25-SEP-2000; 2000US-0235077P.
PR	XX		25-SEP-2000; 2000US-0235082P.
PR	XX		25-SEP-2000; 2000US-0235134P.
PR	XX		25-SEP-2000; 2000US-0235280P.
PR	XX		26-SEP-2000; 2000US-0235637P.
PR	XX		26-SEP-2000; 2000US-0235638P.
PR	XX		27-SEP-2000; 2000US-0235711P.
PR	XX		27-SEP-2000; 2000US-0235720P.
PR	XX		27-SEP-2000; 2000US-0235840P.
PR	XX		27-SEP-2000; 2000US-0235863P.
PR	XX		28-SEP-2000; 2000US-0236028P.
PR	XX		28-SEP-2000; 2000US-0236032P.
PR	XX		28-SEP-2000; 2000US-0236033P.
PR	XX		28-SEP-2000; 2000US-0236034P.
PR	XX		28-SEP-2000; 2000US-0236109P.
PR	XX		28-SEP-2000; 2000US-0236111P.
PR	XX		29-SEP-2000; 2000US-0236642P.
PR	XX		29-SEP-2000; 2000US-0236891P.
PR	XX		02-OCT-2000; 2000US-0237172P.
PR	XX		02-OCT-2000; 2000US-0237173P.
PR	XX		02-OCT-2000; 2000US-0237278P.
PR	XX		02-OCT-2000; 2000US-0237294P.
PR	XX		02-OCT-2000; 2000US-0237295P.
PR	XX		02-OCT-2000; 2000US-0237316P.
PR	XX		03-OCT-2000; 2000US-0237425P.
PR	XX		03-OCT-2000; 2000US-0237598P.
PR	XX		03-OCT-2000; 2000US-0237604P.
PR	XX		03-OCT-2000; 2000US-0237606P.
PR	XX		03-OCT-2000; 2000US-0237608P.
PR	XX		01-NOV-2000; 2000US-0244867P.
PR	XX		01-NOV-2000; 2000US-0245084P.
PA	XX		(AVAL-) AVALON PHARM.
XX	XX		
XX	PI		Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
XX	XX		Soppet DR, Weaver Z;
XX	XX		
XX	XX		WPI; 2002-188264/24.
PT	XX		
PT	XX		Screening for anti-neoplastic agent involves exposing cells to a chemical
PT	XX		agent to be tested for anti-neoplastic activity, and determining a change
PT	XX		in expression of a gene of a signatue gene set.

```

XX  Claim 1; SEQ ID NO 3344; 44pp; English.
XX
CC  The present invention describes a method (M1) for screening for an anti-
CC  neoplastic agent. The method involves exposing cells to a chemical agent
CC  to be tested for anti-neoplastic activity, determining a change in
CC  expression of at least one gene (1) of a signature gene set, where (1)
CC  comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC  to ABL70110), or is at least 95% identical to (S), where a change in
CC  expression is indicative of anti-neoplastic activity. (1) has cytosolic
CC  activity and can be used in gene therapy. M1 can be used for screening an
CC  anti-neoplastic agent, and can be used for producing a product which is
CC  the data collected with respect to the anti-neoplastic agent as a result
CC  of M1, and the data is sufficient to convey the chemical structure and/or
CC  properties of the agent. M1 can be used in the treatment of cancer such
CC  as colon, breast, stomach, lung, thyroid, oesophagel, ovarian, kidney,
CC  prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell,
CC  cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC  cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC  tumour
CC
SQ  Sequence 1421 BP, 419 A, 325 C, 292 G, 385 T, 0 U, 0 Other;

Alignment Scores:
Pred. No.: 1.74e-31 Length: 1421
Score: 294.00 Matches: 55
Percent Similarity: 98.28% Conservative: 2
Best Local Similarity: 94.83% Mismatch: 1
Query Match: 95.77% Indels: 0
DB: Gaps: 0

US-10-031-158-14 (1-58) x ABL65007 (1-1421)

QY 1 MetGlnMetPheProProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
DB 423 ATGCAGATGTTTCCCCCAAGCCACCATATTTTCTTCTTCATATGTCGAAACAAACTCC 483
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 483 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCACGATATTATTAAAGTAC 543
QY 41 ILeGlyLeuLeuArgArgAlaThrArgPheTyrASPProArgArgGlyThrPro 58
DB 543 ATTGGCAAGAAAAGAGAGCAACACGATTCTGGATCTCCAGAGGGGGAACACCA 596

RESULT 24
ADE06987
ID ADE06987 standard; DNA; 1560 BP.
XX
AC ADE06987;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel coding sequence (useful for identifying genetic disorders) #53.
XX
KM novel gene; novel protein; tissue marker; molecular weight marker;
XX  chromosome marker; genetic disorder; gene; ds.
XX
OS Unidentified.
XX
PN MO2003054152-A2.
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002MO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.
XX  11-DEC-2001; 2001US-0339453P.
XX  14-MAR-2002; 2002US-0365091P.
XX  14-MAR-2002; 2002US-0365384P.
XX  12-APR-2002; 2002US-0372381P.
XX  12-APR-2002; 2002US-0372615P.
XX  22-APR-2002; 2002US-00128558.

```

PR 24-APR-2002; 2002US-0376045P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J,
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Dimanac RT, Wang Z,
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ,
XX
XX WPI: 2003-569235/53.
DR P-PSDB; ADB07898.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
XX Claim 1; SEQ ID NO 53; 1177bp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence represents a gene of the
CC invention.
XX
XX Sequence 1560 BP; 444 A; 363 C; 322 G; 431 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.97e-31 Length: 1560
Score: 294.00 Matches: 55
Percent Similarity: 98.28% Conservative: 2
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 95.77% Indels: 0
DB: 10 Gaps: 0
US-10-031-158-14 (1-58) x ADE06987 (1-1560)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 559 ATGCAGATGTTCCCCCAAGCCCACTATTTCCTTCCTCAATGCTGAACAAACATCC 618
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnSerLeuMetLeuLeuArgTyr 40
DB 619 AGAAGGCTGGAACATACCTTCTCTTCTTGAGAAATTTTCCCAATATTATTAAAGATAC 678
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
DB 679 ATTGGCAAGAAAGAGACACAGATTCTGCGATCCAGAGAGGGAACACCA 732
RESULT 25
ABZ35381
ID ABZ35381 standard; cDNA; 1799 BP.
XX
XX AC ABZ35381;
XX
XX 05-FEB-2003 (first entry)
XX
XX Human gene expression profile polynucleotide SEQ ID NO 492.
DE
XX Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW gene expression; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200274979-A2.
XX
XX 26-SEP-2002.
XX
XX 20-MAR-2002; 2002WO-US008456.
XX
XX PF

PR 20-MAR-2001; 2001US-0276947P.
XX
XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
XX Wan J, Wang Y;
PI
XX
XX WPI: 2002-740862/80.
DR
XX
XX New gene expression profile generated from primary, endothelial,
PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer.
XX
XX Example 3; Page 634-635; 850bp; English.
XX
XX The invention relates to a gene expression profile comprising one or more
CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes, DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC identifications and in identifying promising antibiotics, antiviral or
CC antifungal agents
SQ
Sequence 1799 BP; 499 A; 434 C; 396 G; 470 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 2.39e-31 Length: 1799
Score: 294.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 95.77% Indels: 1
DB: 6 Gaps: 0
US-10-031-158-14 (1-58) x ABZ35381 (1-1799)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 845 ATGCAGATGTTCCCCCAAGCCCACTATTTCCTTCCTCAATGCTGAACAAAGCTCC 904
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnSerLeuMetLeuLeuArgTyr 40
DB 905 AGAAGGCTGGAACATACCTTCTCTTCTTGAGAAATTTTCCCGATGTTATTAAAGATAC 964
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
DB 965 ATTGGCAAGAAAGAGACACAGATTCTGCGATCCAGAGAGGGAACACCA 1019
RESULT 26
AAQ37617
ID AAQ37617 standard; cDNA; 1080 BP.
XX
XX AC AAQ37617;
XX
XX 25-MAR-2003 (revised)
XX
XX 21-JUN-1993 (first entry)
XX
XX TCR gamma protein (cell line MOLT-13).
DE

XX T cell antigen receptor; TCR; IDP2; PEER; MOLT-3; immunoglobulin;
KM cell surface; delta chain; Form 2bc; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 37..1008 /*tag= a
XX FT /label= gamma TCR
XX FT /note= "comprises S, V, N, J, C1, C1B, C1C and C1I
XX FT regions"
XX
XX US185250-A.
XX
XX 09-FEB-1993.
XX
XX 13-JAN-1989; 89US-00297661.
XX
XX 03-JUL-1986; 86US-00882100.
XX PR 19-FEB-1987; 87US-00016252.
XX PR 29-OCT-1987; 87US-00115256.
XX PR 29-APR-1988; 88US-00187698.
XX
XX (TCEL-) T CELT SCI INC.
XX PA (DAND) DANA FARBER CANCER INST INC.
XX
XX Brenner MB, Seidman J, Strominger JL, Ip SH, Krangel MS, Band H,
XX WPI; 1993-067162/08.
XX DR F-PSDB; AAR32906.
XX
XX Gamma, delta T-cell antigen receptor heterodimer - used to diagnose
XX PT immune deficiencies e.g. AIDS and stimulate lymphocytes.
XX
XX Disclosure; Fig 10A-B; 99pp; English.
XX
XX Example 11 describes the three forms of the human TCR gamma/delta. The
XX CC new form of the human gamma/delta TCR consists of a 40 kD TCR gamma
XX CC glycoprotein (termed Form 2bc) noncovalently associated with a TCR delta
XX CC chain. Clone M13k corresponds to a full length, in frame, gamma TCR
XX CC transcript. (Updated on 25-MAR-2003 to correct PF field.)
XX
XX SQ Sequence 1080 BP; 324 A; 260 C; 231 G; 265 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4.45e-31 Length: 1080
Score: 290.00 Matches: 54
Percent Similarity: 98.28% Conservative: 3
Best Local Similarity: 93.10% Mismatches: 1
Query Match: 94.46% Indels: 0
DB: 2 Gaps: 0
US-10-031-158-14 (1-58) x AAQ37617 (1-1080)
QY 1 MetGlnMePheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
DB 452 ATGCAGATGTTTCCCAAGCCACATATTTCTTCCTTGATGCTGAAACAAACTCC 511
QY 21 ArgArgLeuGlnIuHsrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTYr 40
DB 512 AGAAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCAGATATTATTAAGATAC 571
QY 41 ILGLYLYELYSARGALATHARGPHEPTASPProARGrglYThrPro 58
DB 572 ATTGGCAAGAAAGAAAGCAACACGATTCGGATCCGAGAGGGAAACACCA 625
RESULT 27
ADE08798
ID ADE08798 standard; DNA; 2658 BP.
XX
XX ADE08798;
XX

DT 29-JAN-2004 (first entry)
XX
XX DE Novel DNA-related contig nucleotide sequence #42.
XX
XX KM novel gene; novel protein; tissue marker; molecular weight marker;
XX KM chromosome marker; genetic disorder; contig; ds.
XX
XX Unidentified.
XX OS
XX PN WO2003054152-A2.
XX
XX 03-JUL-2003.
XX
XX 10-DEC-2002; 2002WO-US039555.
XX PF
XX 11-DEC-2001; 2001US-0339739P.
XX PR 14-MAR-2002; 2001US-0339453P.
XX PR 14-MAR-2002; 2002US-0365091P.
XX PR 14-MAR-2002; 2002US-0365384P.
XX PR 12-APR-2002; 2002US-0372381P.
XX PR 12-APR-2002; 2002US-0372615P.
XX PR 22-APR-2002; 2002US-00128558.
XX PR 24-APR-2002; 2002US-0376045P.
XX
XX (HYSB-) HYSBQ INC.
XX PA
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Dymnac RT, Wang Z;
XX PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX XX WPI; 2003-569235/53.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
XX PT analysis, characterization or therapeutic use, or as markers for tissues
XX PT in which the corresponding protein is preferentially expressed.
XX
XX Disclosure; SEQ ID NO 1864; 1177pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
XX CC proteins. The DNA and protein sequences of the invention are useful as:
XX CC markers for tissues in which the corresponding protein is preferentially
XX CC expressed; as molecular weight markers on gels; as chromosome markers or
XX CC tags; to identify chromosomes or to map related gene positions; and to
XX CC compare with endogenous DNA sequences in patients to identify potential
XX CC genetic disorders. The present DNA sequence was used in the
XX CC exemplification of the invention.
XX
XX SQ Sequence 2658 BP; 745 A; 585 C; 578 G; 750 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.5e-30 Length: 2658
Score: 290.00 Matches: 54
Percent Similarity: 98.28% Conservative: 3
Best Local Similarity: 93.10% Mismatches: 1
Query Match: 94.46% Indels: 0
DB: 10 Gaps: 0
US-10-031-158-14 (1-58) x ADE08798 (1-2658)
QY 1 MetGlnMePheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
DB 497 ATGCAGATGTTTCCCAAGCCACATATTTCTTCCTTGATGCTGAAACAAACTCC 556
QY 21 ArgArgLeuGlnIuHsrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTYr 40
DB 557 AGAAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCAGATATTATTAAGATAC 616
QY 41 ILGLYLYELYSARGALATHARGPHEPTASPProARGrglYThrPro 58
DB 617 ATTGGCAAGAAAGAAAGCAACACGATTCGGATCCGAGAGGGAAACACCA 670
RESULT 28
ADP10449

ID	ADP10449 standard; DNA; 1046 BP.
XX	
AC	ADP10449;
XX	
D7	12-AUG-2004 (first entry)
XX	
DE	Reference mRNA sequences for marker probe #126.
XX	
KW	transplant rejection; immune system; rheumatoid arthritis; lupus; inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
OS	Homo sapiens.
XX	
PN	WO2004042346-A2.
PD	
XX	
PD	21-MAY-2004.
XX	
PF	24-APR-2003; 2003MO-US012946.
XX	
PR	24-APR-2002; 2002US-00131831.
PR	20-DEC-2002; 2002US-00325899.
XX	
PA	(EXPR-) EXPRESSION DIAGNOSTICS INC.
XX	
PI	Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M; Rosenberg S;
XX	
DR	WPI; 2004-400724/37.
PT	
PT	Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic islet, lung, bone marrow or stem cell transplant rejection, in an individual, comprises detecting the expression level of the genes.
PS	
XX	Claim 80; SEQ ID NO 458; 1762bp; English.
CC	The present invention relates to diagnosing or monitoring transplant rejection, e.g. cardiac or kidney transplant rejection, in an individual comprising detecting the expression level of one or more genes. The method, system and kits are useful in diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic islet, lung, bone marrow or stem cell transplant rejection. xenotransplant rejection or mechanical organ replacement rejection, in an individual. The method is also useful in assessing the immune status of an individual. The methods are also useful in diagnosing and monitoring diseases that involve the immune system, e.g. rheumatoid arthritis, lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or viral, bacterial or fungal infection. The present sequence represents a cDNA reference sequence for a 50 mer oligonucleotide marker for diagnosis and monitoring of allograft rejection and other disorders.
SQ	Sequence 1046 BP; 314 A; 238 C; 201 G; 293 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	3, 04e-30 Length: 1046
Score:	284.00 Matches: 53
Percent Similarity:	96.55% Conservative: 3
Best Local Similarity:	91.38% Mismatches: 2
Query Match:	92.51% Indels: 0
DB:	12 Gaps: 0
US-10-031-158-14 (1-58) x ADP10449 (1-1046)	
OY	1 MetGlnMetPhProSerProLeuPhePhePheLeuGlnLeuLysGlnSerSer 20
Db	13 ATGCAGATGTTCCTCCCAAGCCACTATTTCCTTCCTCGATTGGTGAACAACAACTCC 72
OY	21 ATGATGLeuGluHisThrPheValPheLeuArgAsnPhSerSerLeuMetLeuLeuAArgTYr 40
Db	73 AGAAGCGTGGAACATATCTTTGCTCTTCGAGAATAATTTTCCAGATATTATTAAGATAC 132
OY	41 TleglyLVbLVbEArgArgAlaThrArgPheTrpAspProArgArgGLyThrPro 58

DB 133 ATTGGCAGAAAAGACGCAACAGATTCTTGCGATCCAGAGGGGAACACA 186

RESULT 29
ID ABK64529 standard; DNA; 1586 BP.
XX ABK64529,
AC ABK64529,
DT 18-JUN-2002 (first entry)
DE Human benign prostatic hyperplasia gene #424.
KM Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX Homo sapiens.
OS MO200212440-A2.
PN 14-FEB-2002.
PD 07-AUG-2001; 2001WO-US024708.
PE 07-AUG-2000; 2000US-022323P.
PR 05-JUN-2001; 2001US-00873319.
PX (GENE-) GENE LOGIC INC.
PA (NLSB) JAPAN TOBACCO INC.
PI Munger WE, Kulikarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
DR
XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells.

PS Disclosure; Page 247-248; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles. (I)
CC is useful for diagnosing the onset or progression of BPH. (II) is useful
CC for identifying an agent that modulates the onset or progression of BPH.
CC The methods are useful to present information identifying the expression
CC level in a tissue or cells, by comparing the expression level of genes
CC given in the specification in the tissue or cells to the level of
CC expression of gene in the database, and displaying the expression levels
CC of at least one gene in the tissue or cell sample compared to the
CC expression level in BPH. Agents using (II) are useful for treating BPH or
CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic
CC hyperplasia gene sequences of the invention
XX
XX Sequence 1586 BP; 449 A; 369 C; 330 G; 438 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.32e-30 Length: 1586
Score: 284.00 Matches: 53
Percent Similarity: 96.55% Conservative: 3
Best Local Similarity: 91.38% Mismatches: 2
Query Match: 92.51% Indels: 0
DB: 6 Gaps: 0

US-10-031-158-14(1-58) x ABK64529 (1-1586)

```

QY      1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInserSer 20
DB      553 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCCCTTGATGCTGAAACAAACTCC 612
QY      21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB      613 AGAAGGCTGGAACATATCTTGTGCTTCTTGAGAAATTTTCCAGATATTATTAAAGATAC 672
QY      41 IleGlyLysLysArgArgAlaThrArgPheTTPaerProArgArgGlyThrPro 58
DB      673 ATTGGCAAGAAAAGAGACCAACACGATTTCTGGATCCAGAGGGGAAACACCA 726

RESULT 30
AAN91698
ID      AAN91698 standard; DNA; 1080 BP.
XX
XX      AAN91698;
XX
XX      25-MAR-2003 (revised)
DT      31-OCT-2002 (revised)
DT      26-APR-1990 (first entry)
XX
XX      Sequence of MOLT-13 cells gamma T cell antigen receptor (Form 2bc).
DE
XX      Human gamma T cell antigen receptor; gamma TCR; MOLT-13 cells;
XX      clone M13k.
XX
XX      Homo sapiens.
XX
XX      Key CDS
XX      Location/Qualifiers
XX      37..1008
XX      /tag= b
XX      misc_signal 37..46
XX      /tag= a
XX      80..84
XX      /tag= c
XX      /product= "Variable region"
XX      391..394
XX      /tag= d
XX      /product= "N-region"
XX      398..406
XX      /tag= e
XX      /product= "Joining region"
XX      439..448
XX      /tag= f
XX      /product= "CI constant region"
XX      770..778
XX      /tag= g
XX      /product= "CIIB constant region"
XX      817..826
XX      /tag= h
XX      /product= "CIIC constant region"
XX      871..880
XX      /tag= i
XX      /product= "CIII constant region"
XX
XX      misc_RNA
XX      WO890396-A.
XX      PN
XX      PD
XX      05-MAY-1989.
XX
XX      28-OCT-1988; 88MO-US003869.
XX
XX      29-OCT-1987; 87US-00115256.
XX      PR      29-APR-1988; 88US-00187698.
XX
XX      (TCELL-) TCELL SCIENCES INC.
XX      (DAND-) DANA-FABER CANCER INST.
XX      (HARD) HARVARD COLLEGE.
XX
XX      Brenner MB, Ip SH, Seidman J;
XX      PI
XX      WPI; 1989-150856/20.
XX

```

```

DR      P-PSDB; AAP91885.
XX
XX      Human gamma, delta T cell antigen receptor - studied by isolation of
PT      polypeptide(s), nucleic acids and monoclonal antibodies reactive with
PT      epitope(s).
XX
XX      Fig 14; Page 7; 104pp; English.
XX
XX      The CDS (tag b) codes for human gamma T cell antigen receptor polypeptide
CC      termed Form 2bc. The sequence can be DNA or RNA. It was prepd. as
CC      follows: a cDNA library was prepd. from MOLT-13 poly(A)+ RNA in lambda
CC      g10 vector. This library was amplified on the E.coli strain C60 Hfl and
CC      screened. cDNA clone M13k was selected for sequencing. It corresponds to
CC      a full length, in-frame gamma T cell antigen receptor transcript.
CC      (Updated on 31-OCT-2002 to add missing 06 field.) (Updated on 25-MAR-2003
CC      to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      Sequence 1080 BP; 324 A; 257 C; 233 G; 266 T; 0 U; 0 Other;
SQ
XX

Alignment Scores:
Pred. No.: 1.63e-29 Length: 1080
Score: 279.00 Matches: 53
Percent Similarity: 96.55% Conservative: 3
Best Local Similarity: 91.38% Mismatches: 2
Query Match: 90.88% Indels: 0
DB: 1 Gaps: 0

US-10-031-158-14 (1-58) x AAN91698 (1-1080)
QY      1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInserSer 20
DB      452 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCCCTTGATGCTGAAACAAACTCC 511
QY      21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB      512 AGAAGGCTGGAACATATCTTGTGCTTCTTGAGAAATTTTCCAGATATTATTAAAGATAG 571
QY      41 IleGlyLysLysArgArgAlaThrArgPheTTPaerProArgArgGlyThrPro 58
DB      572 ATTGGCAAGAAAAGAGACCAACACGATTTCTGGATCCAGAGGGGAAACACCA 625

RESULT 31
ABV37856
ID      ABV37856 standard; cDNA; 533 BP.
XX
XX      ABV37856;
XX
XX      16-SEP-2002 (first entry)
XX
XX      Human prostate expression marker cDNA 37847.
DE
XX      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX      pharmacogenomic marker; gene; ss.
XX
XX      Homo sapiens.
XX
XX      WO200160860-A2.
XX      PN
XX      PD
XX      23-AUG-2001.
XX
XX      20-FEB-2001; 2001WO-US005171.
XX
XX      17-FEB-2000; 2000US-01831319P.
XX      PR      16-MAR-2000; 2000US-0189862P.
XX      PR      25-MAY-2000; 2000US-0207454P.
XX      PR      09-JUN-2000; 2000US-0211314P.
XX      PR      18-JUL-2000; 2000US-0219007P.
XX      PR      13-DEC-2000; 2000US-0255281P.
XX
XX      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX      PA
XX      PI
XX      Schlegel R, Endege WO, Monahan JE;
XX

```

DR WPI, 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1, Page 7745, 11750bp, English.

XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 533 BP; 180 A; 116 C; 103 G; 132 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	4,38e-27	Length:	533
Score:	259.00	Matches:	54
Percent Similarity:	94.74%	Conservative:	0
Best Local Similarity:	94.74%	Mismatches:	3
Query Match:	84.36%	Indels:	1
DB:	5	Gaps:	0

US-10-031-158-14 (1-58) x ABV37856 (1-533)

OY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSerArg 21

DB 123 CAGAGTGTTCCTC--AGCCACTATNTTCTTCCTCAATGCTGAAACAAAGCTCCAGA 180

OY 22 ArgLeuGlnHisThrPheVal1PheLeuArgAnPheSerLeuMetLeuLeuArgTyr 41

DB 181 AGCGTGGACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATACAT 240

OY 42 GlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 241 GGCAAGAAAGAAAGAGCAACAGCATTCCTGATCCAGAGGGAACACCA 291

RESULT 32

AAC04122
ID AAC04122 standard; cDNA; 460 BP.

XX AAC04122;

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 8197.

XX Human; 5' EST: expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

XX 06-SEP-2000.

PF 21-FEB-2000; 2000EP-00200610.

PR 26-FEB-1999; 99US-0122487P.

XX (GSEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1, SEQ ID NO 8197, 71bp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors

SQ Sequence 460 BP; 162 A; 83 C; 90 G; 121 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	4,93e-26	Length:	460
Score:	251.00	Matches:	56
Percent Similarity:	96.55%	Conservative:	0
Best Local Similarity:	96.55%	Mismatches:	2
Query Match:	81.76%	Indels:	2
DB:	3	Gaps:	0

US-10-031-158-14 (1-58) x AAC04122 (1-460)

OY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20

DB 200 ATGCAAGTGTTCCTCCCAAGCCCACTATTTTCTTCCTCAATGCTGAAACAAAGCTC- 258

OY 21 ArgArgLeuGlnHisThrPheVal1PheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 259 AGAAGCTGGAAACAAACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 318

OY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 319 ATTGGGAARARAAAGAGCAACACGATC-TGGATCCAGAGGGAACACCA 371

RESULT 33

AAF15623
ID AAF15623 standard; cDNA; 1155 BP.

XX AAF15623;

DT 13-MAR-2001 (first entry)

XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:58.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;
KW vulnereary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.

OS Homo sapiens.

PN WO20005174-A1.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US005968.

PR 12-MAR-1999; 99US-0124270P.

XX

XX

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
DR P-PSDB; AAB56420.
XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
XX
PS Claim 1; Page 655-656; 233bp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
XX Sequence 1155 BP; 359 A; 242 C; 225 G; 317 T; 0 U; 12 Other;
SQ
XX
XX Alignment Scores:
Pred. No.: 1,21e-24 Length: 1155
Score: 245.00 Matches: 51
Percent Similarity: 91.07% Conservative: 0
Best Local Similarity: 91.07% Mismatches: 5
Query Match: 79.80% Indels: 1
DB: 3 Gaps: 0
US-10-031-158-14 (1-58) x AAF15623 (1-1155)
QY 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLeuYgInSerSer 20
DB 175 ATGCAAGATGTTTCCCAAGCCCACTAT-TTCTTCTTCTTCTGAAACAAACTCC 233
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 234 AGAAGGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCGATTTTAAAGATAC 293
QY 41 IleGlyLysLysArgArgAlaThrArgPheThrPheProArgArgIly 56
DB 294 ATTGGAGAGAAAGAGAGACACGATTTGCGATCCCAAGAGGGG 341
RESULT 34
ABV29989/c
ID ABV29989 standard; cDNA; 1022 BP.
XX
XX ABV29989;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 29980.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200160860-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US005171.
PF

XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX
PS Claim 1; Page 6477; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC / (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX
SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
XX
XX Alignment Scores:
Pred. No.: 5,29e-24 Length: 1022
Score: 240.00 Matches: 47
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 0
Query Match: 78.18% Indels: 0
DB: 5 Gaps: 0
US-10-031-158-14 (1-58) x ABV29989 (1-1022)
QY 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLeuYgInSerSer 20
DB 589 ATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGAAACAAACTCC 530
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATAC 470
QY 41 IleGlyLysLysArgArgAlaThrArgPheThrPheProArgArgIly 48
DB 469 ATTGGAGAGAAAGAGAGACATCG 446
RESULT 35
ABV24072/c
ID ABV24072 standard; cDNA; 1022 BP.
XX
XX ABV24072;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 24063.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200160860-A2.
PN

XX 23-AUG-2001.
PD
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 4469; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (i) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (i) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5,29e-24 Length: 1022
Score: 240.00 Matches: 47
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 0
Query Match: 78.18% Indels: 0
DB: Gaps: 0
US-10-031-158-14 (1-58) X ABV24072 (1-1022)
QY 1 MetGIMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
DB 589 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCAATGTTGTAACAAAGCTCC 530
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
DB 529 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 470
QY 41 TleGlyLyblybArgArgAlaThr 48
DB 469 ATTGGAGAGAAAGAAAGAGCATCG 446
RESULT 36
ABV24124/C
ID ABV24124 standard; cDNA; 1022 BP.
XX
AC ABV24124;
XX
DT 16-SBP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 24115.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.

XX Homo sapiens.
OS
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 4484-4485; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (i) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (i) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5,29e-24 Length: 1022
Score: 240.00 Matches: 47
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 0
Query Match: 78.18% Indels: 0
DB: Gaps: 0
US-10-031-158-14 (1-58) X ABV24124 (1-1022)
QY 1 MetGIMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
DB 589 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCAATGTTGTAACAAAGCTCC 530
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
DB 529 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 470
QY 41 TleGlyLyblybArgArgAlaThr 48
DB 469 ATTGGAGAGAAAGAAAGAGCATCG 446
RESULT 37
ABV25823/C
ID ABV25823 standard; cDNA; 1022 BP.
XX
AC ABV25823;
XX
DT 16-SBP-2002 (first entry)
XX

DE Human prostate expression marker CDNA 25814.
 XX
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; 88.
 OS Homo sapiens.
 XX
 XX WO200160860-A2.
 XX
 XX 23-AUG-2001.
 XX
 XX 20-FEB-2001; 2001WO-US005171.
 XX
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 XX 25-MAY-2000; 2000US-0207454P.
 XX 09-JUN-2000; 2000US-0211314P.
 XX 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 XX Claim 1; Page 5198; 11750pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC
 XX SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 Pred. No.: 5,29e-24 Length: 1022
 Score: 240.00 Matches: 47
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 97.92% Mismatches: 0
 Query Match: 78.18% Indels: 0
 DB: Gaps: 0
 US-10-031-158-14 (1-58) x ABV25823 (1-1022)
 QY 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInSerSer 20
 DB 589 ATGACGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCGAAACAAAGCTCC 530
 QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 470
 QY 41 IleGlyLylylArgArgAlaThr 48
 DB 469 ATTGGGAAGAAAAGAGCATCG 446
 RESULT 38
 ABV24034/C
 ID ABV24034 standard; CDNA; 1022 BP.
 XX

AC ABV24034;
 XX
 XX 16-SEP-2002 (first entry)
 XX
 XX Human prostate expression marker CDNA 24025.
 DE
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; 88.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200160860-A2.
 XX
 XX 23-AUG-2001.
 XX
 XX 20-FEB-2001; 2001WO-US005171.
 XX
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 XX 25-MAY-2000; 2000US-0207454P.
 XX 09-JUN-2000; 2000US-0211314P.
 XX 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 XX Claim 1; Page 4456; 11750pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC
 XX SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 Pred. No.: 5,29e-24 Length: 1022
 Score: 240.00 Matches: 47
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 97.92% Mismatches: 0
 Query Match: 78.18% Indels: 0
 DB: Gaps: 0
 US-10-031-158-14 (1-58) x ABV24034 (1-1022)
 QY 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInSerSer 20
 DB 589 ATGACGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCGAAACAAAGCTCC 530
 QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 470
 QY 41 IleGlyLylylArgArgAlaThr 48
 DB 469 ATTGGGAAGAAAAGAGCATCG 446


```
RESULT 39
ABV23986/c
ID ABV23986 standard; cDNA; 1022 BP.
XX
XX
AC ABV23986;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 23977.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-018319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4439; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (c) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (d) selecting a composition for inhibiting prostate cancer in a patient;
XX (e) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,296-24 Length: 1022
XX Score: 240.00 Matches: 47
XX Percent Similarity: 100.00% Conservative: 1
XX Best Local Similarity: 97.92% Mismatches: 0
XX Query Match: 78.18% Indels: 0
XX DB: 5 Gaps: 0
XX
XX US-10-031-158-14 (1-58) x ABV23986 (1-1022)
XX
XX 1 MetGlnMetPheProSerPheLeuPhePhePheGlnLeuLeuYsgInSerSer 20
XX
XX 589 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCTCAATGCTGAAACAAAGCTCC 530
XX
XX 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
XX
XX 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 470
XX
```

```
XX
XX 41 IleGlyTylsYArgArgAlaThr 48
XX
XX 469 ATTGGAGAGAAAAGAGAGCATCG 446
XX
XX RESULT 40
XX ABV23991/c
XX ID ABV23991 standard; cDNA; 1022 BP.
XX
XX
XX AC ABV23991;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 23982.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-018319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4442; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (c) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (d) selecting a composition for inhibiting prostate cancer in a patient;
XX (e) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,296-24 Length: 1022
XX Score: 240.00 Matches: 47
XX Percent Similarity: 100.00% Conservative: 1
XX Best Local Similarity: 97.92% Mismatches: 0
XX Query Match: 78.18% Indels: 0
XX DB: 5 Gaps: 0
XX
XX US-10-031-158-14 (1-58) x ABV23991 (1-1022)
XX
XX 1 MetGlnMetPheProSerPheLeuPhePhePheGlnLeuLeuYsgInSerSer 20
XX
XX 589 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCTCAATGCTGAAACAAAGCTCC 530
XX
```

Qy	21	ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr	40
Db	529	AGAAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC	470
Qy	41	ILGLYLYSLYSAArgArgAlaThr	48
Db	469	ATTGGAGAGAAAAAGAAAGCATCG	446

Search completed: December 8, 2004, 08:22:18
Job time : 253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2004, 07:44:01 ; Search time 1517 Seconds

(without alignments)
1393.213 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307
Sequence: 1 MCMPPSPLEFFLQKQSS.....RYIGKKRRATFPDPRRGT 58

Scoring table:

BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3282875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgnt 1/USPRO.spool/US10031158/funat.06122004.082659.15900/app_query.fasta_1.199
-DB=EST -OPMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi -LIST=45
-DOCNALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10031158 @CGN 1.1 2607 @runat.06122004.082659.15900 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g861:*
9: gb_g862:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	427	2	BF679325 602153825
2	307	100.0	510	1	AV716204 AV716204
3	307	100.0	596	1	AV715641 AV715641
4	307	100.0	603	1	AV714189 AV714189
5	307	100.0	803	1	AV756492 AV756492
6	307	100.0	818	4	BI771070 603059635
7	307	100.0	821	2	BF679123 602153390
8	307	100.0	861	2	BF681385 602156579
9	307	100.0	865	1	AI557112 PT2.1_13_

10	307	100.0	881	2	BF677648	BF677648 602085529
11	307	100.0	902	2	BF674593	BF674593 602137811
12	307	100.0	927	3	BF679476	BF679476 602153280
13	307	100.0	1058	4	BM544213	BM544213 AGENCOURT
14	307	100.0	3533	3	BC030554	BC030554 Homo sapi
15	304	99.0	789	5	BU199143	BU199143 DCBCLA08
16	294	95.8	653	3	BE974171	BE974171 601680522
17	294	95.8	1005	4	BM919380	BM919380 AGENCOURT
18	294	95.8	1757	3	BC039116	BC039116 Homo sapi
19	290	94.5	434	2	BF747659	BF747659 OVE-BT061
20	290	94.5	477	2	AM087486	AM087486 XB57C01.X
21	290	94.5	681	1	AV714756	AV714756 AV714756
22	290	94.5	1510	3	BC017861	BC017861 Homo sapi
23	289	94.1	1055	2	BF964904	BF964904 602268596
24	288	93.8	523	2	BF675605	BF675605 602083495
25	284	92.5	362	1	AI758378	AI758378 tV68C12.X
26	278	90.6	719	2	BF681238	BF681238 602155502
27	270	87.9	854	2	BF681029	BF681029 602155433
28	268	87.3	629	2	BF679165	BF679165 602153737
29	267	87.0	803	2	BF680840	BF680840 602156088
30	262	85.3	772	4	BI836020	BI836020 603084004
31	259	84.4	808	2	BF673289	BF673289 602136139
32	253	82.4	857	2	BF674457	BF674457 602137231
33	251	81.8	849	2	BF679284	BF679284 602153477
34	248	80.8	995	7	CO647648	CO647648 ILLUMIGEN
35	245	79.8	663	6	CB553506	CB553506 MMSPP004
36	227	73.9	870	2	BF678129	BF678129 602085181
37	226	73.6	637	6	CB551454	CB551454 MMSPP0013
38	224	72.0	670	2	BF677916	BF677916 602084736
39	221	72.0	827	4	BG535408	BG535408 602563069
40	209	68.1	572	2	BF370026	BF370026 MR3-FN000
41	204.5	66.6	864	4	BI834993	BI834993 603088885
42	203	66.1	552	4	BP164784	BP164784 BP164784
43	200	65.1	572	5	BP162849	BP162849 BP162849
44	200	65.1	888	5	BP162346	BP162346 BP162346
45	193	62.9	436	5	BP165939	BP165939 BP165939

ALIGNMENTS

RESULT 1
LOCUS BF679325
DEFINITION 602153825F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294566 5',
mRNA sequence.

ACCESSION BF679325
VERSION BF679325.1 GI:11953220

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 427)

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

TISSUE Procurement: CLONETECH Laboratories, Inc.
CDNA Library Preparation: CLONETECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

plate: L1CM1144 row: b column: 07
High quality sequence stop: 427.

FEATURES
source location/Qualifiers
1..427

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294566"

```

/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NH1_MGC_83"
/notes="Organ: probate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcattatggcc); 5' and 3' adaptor were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGACC-3'
and 3' adaptor sequence:
5'-ATCTAGAGCGCCAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."
```

ORIGIN

Alignment Scores:

Pred. No.:	1,17e-30	Length:	427
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679325 (1-427)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
AV716204
LOCUS AV716204 DCB Homo sapiens cDNA clone DCBAM08 5', mRNA sequence.
DEFINITION AV716204
ACCESSION AV716204.1 GI:10797721
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
COMMENT
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 510)
Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
Lu,G., Cheng,Z. and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source

1..510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBAM08"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"

ORIGIN

/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"

Alignment Scores:

Pred. No.:	1,45e-30	Length:	510
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x AV716204 (1-510)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
AV71641
LOCUS AV71641 DCB Homo sapiens cDNA clone DCBBS05 5', mRNA sequence.
DEFINITION AV71641
ACCESSION AV71641.1 GI:10797158
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
COMMENT
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
Lu,G., Cheng,Z. and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source

1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBBS05"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"

ORIGIN

Alignment Scores:

Pred. No.:	1,74e-30	Length:	596
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x AV71641 (1-596)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
 DB 93 ATCAAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC 152
 QY 21 ArgArgLeuGlnIuIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 153 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 212
 QY 41 IleglYlYlYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 DB 213 ATTGGCAAGAAAGAGAGACACACATCTTGATCTCCAGAGGGGAGAACACCA 266
 RESULT 4
 AV714189 603 bp mRNA linear EST 11-OCT-2000
 LOCUS AV714189 DCB Homo sapiens cDNA clone DCB0E09 5', mRNA sequence.
 DEFINITION AV714189
 ACCESSION AV714189
 VERSION AV714189.1 GI:10795706
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 603)
 AUTHORS Xu,X., Gu,J., Liu,F., Ou,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z. and Han,Z.
 TITLE Homo sapiens cDNA DCB clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 source
 1. 603
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DCB0E09"
 /cell_type="dendritic cells"
 /dev_stage="mature"
 /lab_host="BM25.8"
 /clone_lib="DCB"
 /note="Vector: pTripleX2; Site_1: sfiIA; Site_2: sfiIB"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.77e-30 Length: 603
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-10-031-158-14 (1-58) x AV714189 (1-603)
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
 DB 252 ATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC 311
 QY 21 ArgArgLeuGlnIuIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 312 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 371
 QY 41 IleglYlYlYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 DB 372 ATTGGCAAGAAAGAGAGACACACATCTTGATCTCCAGAGGGGAGAACACCA 425

RESULT 5
 AV756492 803 bp mRNA linear EST 19-OCT-2000
 LOCUS AV756492 BM Homo sapiens cDNA clone BMFBAF12 5', mRNA sequence.
 DEFINITION AV756492
 ACCESSION AV756492
 VERSION AV756492.1 GI:10914340
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 803)
 AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Ou,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
 TITLE Homo sapiens cDNA BM clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 source
 1. 803
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BMFBAF12"
 /tissue="Bone marrow"
 /cell_type="CD34+ hematopoietic stem/progenitor cell"
 /lab_host="BM25.8"
 /clone_lib="BM"
 /note="Vector: pTripleX2; Site_1: sfiIA; Site_2: sfiIB"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.48e-30 Length: 803
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-10-031-158-14 (1-58) x AV756492 (1-803)
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
 DB 434 ATCAAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAAGCTCC 493
 QY 21 ArgArgLeuGlnIuIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 494 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 553
 QY 41 IleglYlYlYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 DB 554 ATTGGCAAGAAAGAGAGACACACATCTTGATCTCCAGAGGGGAGAACACCA 607
 RESULT 6
 BI771070 818 bp mRNA linear EST 25-SEP-2001
 LOCUS BI771070
 DEFINITION 603059635F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208945 5', mRNA sequence.
 ACCESSION BI771070
 VERSION BI771070.1 GI:15762648
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 818)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsbds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LHAM1524 row: e column: 10
High quality sequence stop: 816.
Location/Qualifiers
1. 818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5208945"
/lab_host="DH10B"
/clone_id="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 2,54e-30 Length: 818
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 4
US-10-031-158-14 (1-58) x B1771070 (1-818)
Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 548 ATGCAGAGTGTTCCTCCCAAGCCCACTATTTCTTCTCAATTGCTGAACAAAGCTCC 607
Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 608 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTTAAAGATAC 667
Qy 41 IlleglylsylsArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58
Db 668 ATTGGCAAGAAAAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 721
RESULT 7
LOCUS BF679123 821 bp mRNA linear EST 21-DEC-2000
DEFINITION 602153390P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294247 5',
ACCESSION BF679123
VERSION BF679123.1 GI:11953018
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
COMMENT

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsbds-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LHCN1143 row: d column: 24
High quality sequence stop: 620.
Location/Qualifiers
1. 821
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294247"
/lab_host="DH10B (T1 phage-resistant)"
/clone_id="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and 3' adaptor were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAAGCGCCGAGCGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN
Alignment Scores:
Pred. No.: 2,55e-30 Length: 821
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 2
US-10-031-158-14 (1-58) x BF679123 (1-821)
Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
Db 64 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCC 123
Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 124 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTTAAAGATAC 183
Qy 41 IlleglylsylsArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58
Db 184 ATTGGCAAGAAAAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 237
RESULT 8
LOCUS BF681385 861 bp mRNA linear EST 21-DEC-2000
DEFINITION 602156579P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4297555 5',
ACCESSION BF681385
VERSION BF681385.1 GI:11955280
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCM151 row: n column: 20
 High quality sequence stop: 672.

FEATURES

source

Location/Qualifiers

1. .861
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:429755"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NIH_MGC_83"
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
 site_1: SfiI (ggcgctggcc); site_2: SfiI
 (ggcgctggcc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3'
 and 3' adaptor sequence:
 5'-ATCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA)."

ORIGIN

Alignment Scores:

Pred. No.: 2.7e-30 Length: 861
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-031-158-14 (1-58) x BF681385 (1-861)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
 Db 69 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCATTCGCGAAACAAAGCTCC 128
 Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40
 Db 129 AGAAGCGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 188
 Qy 41 TlEGlYlyslYsArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 189 ATTGCGAAGAAAGAGAGCAACGATTCGGATCCCGAGAGGGAGCAACCA 242

RESULT 9

AI557112 865 bp mRNA linear EST 09-AUG-1999
 LOCUS PT2.1.13.F05.r tumor2 Homo sapiens CDNA 3', mRNA sequence.
 AI557112
 ACCESSION
 VERSION
 AI557112.1 GI:4489475
 EST.
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 Huang,G.M., Ng,W.L., Parkes,J., He,L., Liang,H.A., Gordon,D., Yu,J.
 and Hood,L.
 Prostate cancer expression profiling by cDNA sequencing analysis

JOURNAL
 MEDLINE
 PUBMED
 Genomics 59 (2), 178-186 (1999)
 10409429
 COMMENT
 Leroy Hood

University of Washington
 Department of Molecular Biotechnology, Box 357730, University of
 Washington, Seattle, WA 98195
 Tel: 5106280100
 Fax: 5106280108
 Email: huangsm@yahoo.com

FEATURES

source

Location/Qualifiers

1. .865
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="tumor2"
 /note="Organ: Prostate; Vector: pBluescript; Directional
 CDNA library was constructed using Lambda ZP II kit
 (Stratagene). mRNA was extracted from a frozen prostate
 tumor tissue (Mayo Clinics)."

ORIGIN

Alignment Scores:

Pred. No.: 2.71e-30 Length: 865
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-031-158-14 (1-58) x AI557112 (1-865)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
 Db 59 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCATTCGCGAAACAAAGCTCC 118
 Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40
 Db 119 AGAAGCGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 178
 Qy 41 TlEGlYlyslYsArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 179 ATTGCGAAGAAAGAGAGCAACGATTCGGATCCCGAGAGGGAGCAACCA 232

RESULT 10

BF677648 881 bp mRNA linear EST 21-DEC-2000
 LOCUS 602085529F1 NIH_MGC_83 Homo sapiens CDNA clone IMAGE:4249807 5',
 DEFINITION
 mRNA sequence.

BF677648
 BF677648.1 GI:11951543
 VERSION
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS
 NIH-MGC <http://imgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCM1070 row: 1 column: 08
 High quality sequence stop: 601.

FEATURES

source

Location/Qualifiers

1. .881
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4249807"

/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech) ;
Site 1: SfiI (ggcgctcgcc) ; Site 2: SfiI
(ggcgctcgcc) ; 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.77e-30	Length:	881
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF677648 (1-881)

Qy 1 MeGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYglnSerSer 20
Db 64 ATGCAGATGTTCCCGCAAGCCCACTATTTCTTCTTCATTGCTGAACAAAGCTCC 123

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40
Db 124 AGAAGGCTGGAACATACCTTGCTTCTTGAGAAATTTTCCCGATGTATTAAGATAC 183

Qy 41 IleglylylsySArGAAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 184 ATTGGGAAGAAAGAAAGCAACAGATTTCTGGATCCAGAGGGGAACACCA 237

RESULT 11
LOCUS BF674593 902 bp mRNA linear EST 21-DEC-2000
DEFINITION 60213781P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4274283 5',
mRNA sequence.
ACCESSION BF674593
VERSION BF674593.1 GI:11948488
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS 1 (bases 1 to 902)
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b6-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM1091 row: e column: 04
High quality sequence stop: 629.
Location/Qualifiers
1..902
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4274283"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech) ;
Site 1: SfiI (ggcgctcgcc) ; Site 2: SfiI
(ggcgctcgcc) ; 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.77e-30	Length:	881
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF677648 (1-881)

Qy 1 MeGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYglnSerSer 20
Db 64 ATGCAGATGTTCCCGCAAGCCCACTATTTCTTCTTCATTGCTGAACAAAGCTCC 123

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40
Db 124 AGAAGGCTGGAACATACCTTGCTTCTTGAGAAATTTTCCCGATGTATTAAGATAC 183

Qy 41 IleglylylsySArGAAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 184 ATTGGGAAGAAAGAAAGCAACAGATTTCTGGATCCAGAGGGGAACACCA 237

RESULT 11
LOCUS BF674593 902 bp mRNA linear EST 21-DEC-2000
DEFINITION 60213781P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4274283 5',
mRNA sequence.
ACCESSION BF674593
VERSION BF674593.1 GI:11948488
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS 1 (bases 1 to 902)
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b6-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM1091 row: e column: 04
High quality sequence stop: 629.
Location/Qualifiers
1..902
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4274283"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech) ;
Site 1: SfiI (ggcgctcgcc) ; Site 2: SfiI
(ggcgctcgcc) ; 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

Site 1: SfiI (ggcgctcgcc) ; Site 2: SfiI
(ggcgctcgcc) ; 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.85e-30	Length:	902
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF674593 (1-902)

Qy 1 MeGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYglnSerSer 20
Db 64 ATGCAGATGTTCCCGCAAGCCCACTATTTCTTCTTCATTGCTGAACAAAGCTCC 123

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40
Db 124 AGAAGGCTGGAACATACCTTGCTTCTTGAGAAATTTTCCCGATGTATTAAGATAC 183

Qy 41 IleglylylsySArGAAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 184 ATTGGGAAGAAAGAAAGCAACAGATTTCTGGATCCAGAGGGGAACACCA 237

RESULT 12
LOCUS BF679476 927 bp mRNA linear EST 21-DEC-2000
DEFINITION 602153280P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294507 5',
mRNA sequence.
ACCESSION BF679476
VERSION BF679476.1 GI:11953360
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS 1 (bases 1 to 927)
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b6-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM1143 row: o column: 20
High quality sequence stop: 573.
Location/Qualifiers
1..927
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294507"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech) ;
Site 1: SfiI (ggcgctcgcc) ; Site 2: SfiI
(ggcgctcgcc) ; 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.85e-30	Length:	902
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF674593 (1-902)

Qy 1 MeGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYglnSerSer 20
Db 64 ATGCAGATGTTCCCGCAAGCCCACTATTTCTTCTTCATTGCTGAACAAAGCTCC 123

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40
Db 124 AGAAGGCTGGAACATACCTTGCTTCTTGAGAAATTTTCCCGATGTATTAAGATAC 183

Qy 41 IleglylylsySArGAAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 184 ATTGGGAAGAAAGAAAGCAACAGATTTCTGGATCCAGAGGGGAACACCA 237

RESULT 12
LOCUS BF679476 927 bp mRNA linear EST 21-DEC-2000
DEFINITION 602153280P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294507 5',
mRNA sequence.
ACCESSION BF679476
VERSION BF679476.1 GI:11953360
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS 1 (bases 1 to 927)
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b6-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM1143 row: o column: 20
High quality sequence stop: 573.
Location/Qualifiers
1..927
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294507"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech) ;
Site 1: SfiI (ggcgctcgcc) ; Site 2: SfiI
(ggcgctcgcc) ; 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

and 3' adaptor sequence:
 5'-ATTCTAGAGCGGCGGCGGACATG-dr(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.94e-30	Length:	927
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679476 (1-927)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
 Db 67 ATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAGCTCC 126

Qy 21 ArgArgLeuGlnIshThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 Db 127 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTATTAAAGTAC 186

Qy 41 IIGLYLSYLSARGALAThArgPheTTPAspProArgArgIYhPro 58
 Db 187 ATTGGCAGAAAGAGAGACAGATTTCTGGATCCAGAGGGAGACCA 240

RESULT 13
 BMS44213 1058 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT_6490556 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587705
 DEFINITION 5', mRNA sequence.
 ACCESSION BMS44213
 VERSION BMS44213.1 GI:18775279
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1058)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Invitrogen
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 Distribution by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 plate: LLM12357 row: k column: 02
 High quality sequence start: 12
 High quality sequence stop: 768.
 Location/Qualifiers

FEATURES

source

1..1058
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5587705"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 125"
 /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
 Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 2.1 kb, insert size range 1-3.5 kb. Library is

normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."

ORIGIN

Alignment Scores:

Pred. No.:	3.44e-30	Length:	1058
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x BMS44213 (1-1058)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
 Db 421 ATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAGCTCC 480

Qy 21 ArgArgLeuGlnIshThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 Db 481 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTATTAAAGTAC 540

Qy 41 IIGLYLSYLSARGALAThArgPheTTPAspProArgArgIYhPro 58
 Db 541 ATTGGCAGAAAGAGAGACAGATTTCTGGATCCAGAGGGAGACCA 594

RESULT 14
 BC030554 3533 bp mRNA linear HTC 20-MAY-2002
 LOCUS Homo sapiens, similar to T cell receptor gamma locus, clone
 DEFINITION IMAGE:5213435, mRNA.
 ACCESSION BC030554
 VERSION BC030554.1 GI:20988582
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3533)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (07-MAY-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@ngri.nih.gov
 Ahner, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Laric, P., Legaaspi, R., Maduro, O.L.,
 Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J.,
 Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C.,
 Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 64 Row: k Column: 18
 This clone has the following problem: frame shifted.

FEATURES

source

1..3533
 /organism="Homo sapiens"

US-10-031-158-14 (1-58) x BB974171 (1-653)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
 Db 310 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTTCATTTGTAACAAAGCTCC 369
 Qy 21 ArgArgLeuGlnIshThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40
 Db 370 AGAAGCTGGAACACTACTTCTGTTCTTGAGAAATTTTCCCGATTTATTAAAGTAC 429

Qy 41 Ile-GlyValysArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
 Db 430 ATTGGCAAGAAAGAGAGACACATTTCTGGATCCCGAGGGAACACCA 484

RESULT 17

BM919380 1005 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT 6715666 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748632

DEFINITION 5', mRNA sequence.

ACCESSION BM919380
 VERSION BM919380.1 GI:19369759

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1005)
 NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1AM1276 Row: 1 Column: 09
 High quality sequence stop: 609.

Location/Qualifiers
 1. 1005

FEATURES
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5748632"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 120"
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.76e-28 Length: 1005
 Score: 294.00 Matches: 55
 Percent Similarity: 98.28% Conservative: 2
 Best Local Similarity: 94.83% Mismatches: 1
 Query Match: 95.77% Indels: 0
 DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x BM919380 (1-1005)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
 Db 397 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTTCATTTGTAACAAAGCTCC 456

Qy 21 ArgArgLeuGlnIshThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40
 Db 457 AGAAGCTGGAACACTACTTCTGTTCTTGAGAAATTTTCCCGATTTATTAAAGTAC 516

Qy 41 IleGlyValysArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
 Db 517 ATTGGCAAGAAAGAGAGACACATTTCTGGATCCCGAGGGAACACCA 570

RESULT 18

BC039116 1757 bp mRNA linear HTC 04-MAR-2003
 LOCUS Homo sapiens, similar to T-cell receptor gamma, constant 2, clone
 IMAGE:4829750, mRNA.

ACCESSION BC039116 GI:25058606

VERSION BC039116.1

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1757)
 Strausberg, R.

AUTHORS Direct Submission
 TITLE Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk

EMAIL: cgabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshuyuki and Piero Carninci (RIKEN)

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome

Center code: BCM-HQSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/

Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Wuzny, D.M., Nanavali, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov

Series: IRAP Plate: 72 Row: F Column: 5
 This clone has the following problem: retained intron.

Location/Qualifiers

1. 1757
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4829750"
 /tissue_type="Testis"
 /clone_lib="NIH MGC_97"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Alignment Scores:

Pred. No.: 3.41e-28 Length: 1757
 Score: 294.00 Matches: 55
 Percent Similarity: 98.28% Conservative: 2
 Best Local Similarity: 94.83% Mismatches: 1
 Query Match: 95.77% Indels: 0
 DB: 3 Gaps: 0

US-10-031-158-14 (1-58) x BC039116 (1-1757)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
 Db 700 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTTCATTTGTAACAAAGCTCC 759

Qy 21 ArgArgLeuGluH1sthrPheVal1PheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 Db 760 AGAAGGCTGGAACATCTTGTCTTGTGGAATTTTCCAGATATATTAAGATAC 819
 Qy 41 11gg1ylyslYsAArgAla1sthrArgPheTrpAspProAlaGArgGlyThrPro 58
 Db 820 ATTGGCAGAAAAGAGAGCAACGATTCTGGATCCAGAGGGGAGACCA 873

RESULT 19
 BF747659/c 434 bp mRNA linear EST 10-JAN-2001
 DEFINITION QV2-BT0616-051200-521-c07 BT0616 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF747659
 VERSION BF747659.1 GI:12074335
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL MEDLINE 10737800
 PUBMED 20202683

TITLE Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPSP/ILCR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2et2-QV2-BT0616-
 051200-521-c07&cl=2000-12-05&cl=4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 320.
 Location/Qualifiers

FEATURES
 source
 1..434
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BT0616"
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,226-28 Length: 434
 Score: 290.00 Matches: 54
 Percent Similarity: 98.28% Conservative: 3
 Best Local Similarity: 93.10% Mismatches: 1
 Query Match: 94.46% Indels: 0
 DB: 2 Gaps: 0
 US-10-031-158-14 (1-58) x BF747659 (1-434)

Qy 1 MetGlnMetPheProPsoSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
 Db 305 ATGCAGATGTTTCCCCCAACCCCACTATTTTCTTCTTCGATTCGAAACAAACTCC 246

Qy 21 ArgArgLeuGluH1sthrPheVal1PheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 Db 245 AGAAGGCTGGAACATCTTGTCTTGTGGAATTTTCCAGATATATTAAGATAC 186

Qy 41 11gg1ylyslYsAArgAla1sthrArgPheTrpAspProAlaGArgGlyThrPro 58
 Db 185 ATTGGCAGAAAAGAGAGCAACGATTCTGGATCCAGAGGGGAGACCA 132

RESULT 20
 AM087486/c 497 bp mRNA linear EST 15-OCT-1999
 LOCUS XB67C01.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:2581344.3, similar to gb:U13321 T-CELL RECEPTOR GAMMA CHAIN C
 REGION (HUMAN), mRNA sequence.
 ACCESSION AM087486
 VERSION AM087486.1 GI:6043385
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 1 (bases 1 to 497)
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 TITLE Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-rt@mail.nih.gov
 This clone is available royalty-free through LINT; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 415.
 Location/Qualifiers

FEATURES
 source
 1..497
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2581344"
 /lab_host="DH10B"
 /clone_lib="Soares NFL T GBC S1"
 /note="Organ: pooled; Vector: pRT73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung KBHL19W, testis NRT, and B-cell
 NCI-CGAP-GCB1) were mixed, and 88 clones were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The diver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 1.M.A.G.B. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,66-28 Length: 497
 Score: 290.00 Matches: 54
 Percent Similarity: 98.28% Conservative: 3
 Best Local Similarity: 93.10% Mismatches: 1
 Query Match: 94.46% Indels: 0
 DB: 2 Gaps: 0
 US-10-031-158-14 (1-58) x AM087486 (1-497)

Qy 1 MetGlnMetPheProPsoSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
 Db 373 ATGCAGATGTTTCCCCCAACCCCACTATTTTCTTCTTCGATTCGAAACAAACTCC 314

Qy 21 ArgArgLeuGluH1sthrPheVal1PheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

```

Db      313 AGAAGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCAGATATTTAAGATAC 254
Qy      41 11eg1ylyslvsaargalathrarphettraspProaGATGlyThrPro 58
Db      253 ATTGGCAAGAAAGAGAGGACACGATTCGGATTCAGAGAGGAGAACCA 200

RESULT 21
AV714756      681 bp      mRNA      linear      EST 11-OCT-2000
DEFINITION    AV714756 DCB Homo sapiens cDNA clone DCBBC809 5', mRNA sequence.
ACCESSION     AV714756
VERSION       AV714756.1 GI:10796273
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 681)
AUTHORS       Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
              Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
              Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,D., Fu,G., Ren,S., Zhong,M.,
              Lu,G., Cheng,Z. and Han,Z.
              Homo sapiens cDNA DCB clones
              Unpublished (2000)
              Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn
              This clone is available at CHGC in Shanghai.

FEATURES
  source
    1..681
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="DCBBC809"
    /cell_type="dendritic cells"
    /dev_stage="mature"
    /lab_host="BM25.8"
    /clone_lib="DCB"
    /note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"

ORIGIN
Alignment Scores:
Pred. No.:      3,78e-28      Length:      681
Score:          290.00      Matches:      54
Percent Similarity: 98.28%      Conservative: 3
Best Local Similarity: 93.10%      Mismatches: 1
Query Match:    94.46%      Indels:      0
DB:             1          Gaps:          0

US-10-031-158-14 (1-58) x AV714756 (1-681)

Qy      1 MetGlnMetPheProPseProleuPhePheLeuGlnLeuLeuGlnSerSer 20
Db      179 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCTTCGATTCGAAACAAACTCC 238
Qy      21 ArgArgLeuGlnIsthRpheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyr 40
Db      239 AGAAGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCAGATATTTAAGATAC 298
Qy      41 11eg1ylyslvsaargalathrarphettraspProaGATGlyThrPro 58
Db      299 ATTGGCAAGAAAGAGAGGACACGATTCGGATTCAGAGAGGAGAACCA 352

RESULT 22
LOCUS        BC017861      1510 bp      mRNA      linear      HTC 06-DEC-2001
DEFINITION    Homo sapiens, Similar to T cell receptor gamma locus, clone
IMAGE:4687960, mRNA.

```

```

ACCESSION     BC017861
VERSION       BC017861.1 GI:17389678
KEYWORDS      HTC.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 1510)
AUTHORS       Strausberg,R.
TITLE         Direct Submission
JOURNAL
REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgabs-remail.nih.gov
              Tissue Procurement: CLONTECH
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              DNA Sequencing by: Sequencing Group at the Stanford Human Genome
              Center, Stanford University School of Medicine, Stanford, CA 94305
              Web site: http://www-shgc.stanford.edu
              Contact: (Dickson, Mark) mcd@paxil.stanford.edu
              Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
              R. M.

FEATURES
  source
    1..1510
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4687960"
    /tissue_type="lung"
    /clone_lib="NIH_MGC_77"
    /lab_host="DH10B"
    /note="Vector: pDNR-LIB"

ORIGIN
Alignment Scores:
Pred. No.:      9.73e-28      Length:      1510
Score:          290.00      Matches:      54
Percent Similarity: 98.28%      Conservative: 3
Best Local Similarity: 93.10%      Mismatches: 1
Query Match:    94.46%      Indels:      0
DB:             3          Gaps:          0

US-10-031-158-14 (1-58) x BC017861 (1-1510)

Qy      1 MetGlnMetPheProPseProleuPhePheLeuGlnLeuLeuGlnSerSer 20
Db      479 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCTTCGATTCGAAACAAACTCC 538
Qy      21 ArgArgLeuGlnIsthRpheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyr 40
Db      539 AGAAGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCAGATATTTAAGATAC 598
Qy      41 11eg1ylyslvsaargalathrarphettraspProaGATGlyThrPro 58
Db      599 ATTGGCAAGAAAGAGAGGACACGATTCGGATTCAGAGAGGAGAACCA 652

RESULT 23
LOCUS        BP964904      1055 bp      mRNA      linear      EST 22-JAN-2001
DEFINITION    602268596F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4357102 5',
              mRNA sequence.
ACCESSION     BP964904
VERSION       BP964904.1 GI:12332119
KEYWORDS      EST.

```

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 523)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone sequencing by: Incyte Genomics, Inc.
Gene distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LICM065 row: m column: 23
High quality sequence stop: 510.

FEATURES
Source
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:4247998"
/lab_host="DH10B (TI phase-resistant)"
/clone_1lb="NH MGC 83"
/note="Organ: prostate; Vector: PDNR-LIB (Clontech); Site_1: SfiI (ggcgccctggcc), Site_2: SfiI (ggcattatgccc), 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCGCATTATGGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGCGGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
Alignment Scores:
Pred. NO.: 5.11e-28 Length: 523
Score: 288.00 Matches: 57
Percent Similarity: 96.61% Conservative: 0
Best local Similarity: 96.61% Mismatches: 1
Query Match: 93.81% Indels: 1
Dbs: 2 Gaps: 0

US-10-031-158-14 (1-58) x BF675605 (1-523)

Cy	1	MecGlmetPheProPsoSerProLeuPhepheLeuLeuLeuVegLseSer	20
Dd	64	ATGCAGAGTGTTCCCCCAAGCCACTATTTTCCTTCATATGCTGAACAAGCTCC	123
Cy	21	ArgArgeLuengiUhiStrPh-Phe-ValPheLeuAraAsnPhSeSerLeuWetLeuLeuAraTy	40
Dd	124	AGAAGGCTGGAACATACTGTGTCTTCTTGAGAAATTTTTCCCTGATGTAATTAAGTA	183
Cy	40	TiIGlYLyLyLyLyArGaLaAltnrArPheTyrAspProArGaRgaLyThrPro	58
Dd	184	CATTGGCAAGAAAGAAAGAACACGATTCCTGGGATCCCGAGGAGGGAACACCA	238

RESULT 25
LOCUS AT158378/c 362 bp mrna linear EST 07-MAR-2000
DEFINITION ty68c12.x1 NCI CGAP Kld11 Homo sapiens cDNA IMAGE:2284246 3'
similar to gb:M13231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN)',,
RNA sequence.

ACCESSION AT158378
VERSION AT158378.1 GI:5152101
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 362)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapdb-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 774 Std Error: 0.00
 Seq primer: -400P from Gldco
 High quality sequence stop: 286.
FEATURES
 source
 1..362
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2284246"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Kid11"
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid3 was prepared, and 88 circles were made in vitro. Following NHP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.13e-27 Length: 362
 Score: 284.00 Matches: 53
 Percent Similarity: 96.55% Conservative: 3
 Best Local Similarity: 91.38% Mismatches: 2
 Query Match: 92.51% Indels: 0
 DB: 1 Gaps: 0

US-10-031-158-14 (1-58) x A1758378 (1-362)

Qy 1 MetGlnMetPheProPProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
 Db 349 ATGCAGATGTTTCCCCCAGCCCACTATTATCTTCTTCGATTCGTAACAAACCTCC 290

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
 Db 289 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCACAGATTTATTAAGATAC 230

Qy 41 IlleglylsylsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 229 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCAGAGAGGGAACACCA 176

RESULT 26
 BPe81238 719 bp mRNA linear EST 21-DEC-2000
 LOCUS BPe81238 602155502P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4296504 5', mRNA sequence.
 ACCESSION BPe81238
 VERSION BPe81238.1 GI:11955133
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 719)
AUTHORS NIH-MGC <http://mgc.ncbi.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapdb-remail.nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1CWM149 row: c column: 01
 High quality sequence stop: 673.
FEATURES
 source
 1..719
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4296504"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_83"
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcctcgcc); Site 2: SfiI (ggcctatcgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor sequence: 5'-ATCTTAGAGCCGAGGCGCCGACATC-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.61e-26 Length: 719
 Score: 278.00 Matches: 57
 Percent Similarity: 98.28% Conservative: 0
 Best Local Similarity: 98.28% Mismatches: 1
 Query Match: 90.55% Indels: 1
 DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x BPe81238 (1-719)

Qy 1 MetGlnMetPheProPProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
 Db 68 ATGCAGATGTTTCCCCCAGCCCACTATTATCTTCTTCGATTCGTAACAAAGCTCC 127

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
 Db 128 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCACAGATTTATTAAGATAC 186

Qy 41 IlleglylsylsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
 Db 187 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCAGAGAGGGAACACCA 240

RESULT 27
 BPe81029 854 bp mRNA linear EST 21-DEC-2000
 LOCUS BPe81029 602155433P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4296104 5', mRNA sequence.
 ACCESSION BPe81029
 VERSION BPe81029.1 GI:11954924
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 854)
AUTHORS NIH-MGC <http://mgc.ncbi.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgarbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
CNA Library Preparation: CLONETECH Laboratories, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Plate: L1CM1148 row: b column: 09
High quality sequence stop: 518.
Location/Qualifiers

FEATURES

source

```
1. 854
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4296104"
/lab_host="DH10B (TI phage-resistant)"
/clone_id="NIH_MGC_83"
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech); Site:1: Still (ggccgcctgcgc); Site:2: Still (ggccatcagcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
```

ORIGIN

Alignment Scores:

Pred. No.:	2.3e-25	Length:	854
Score:	270.00	Matches:	53
Percent Similarity:	93.22%	Conservative:	2
Best Local Similarity:	89.83%	Mismatches:	3
Query Match:	87.95%	Indels:	1
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF681029 (1-854)

QY 1 MetGlnMetPheProPheSerProLeuPhePhePheGlnLeuLeuGlnSerSer 20
DB 65 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCTTCAATGCTAAACAAAGCTCC 124

QY 21 ArgArgLeuGluHisThr-PheValPheLeuArganPheSerLeuMetLeuLeuArgTy 40
DB 125 AGAAGGCTGGAACATACCTTTGTTCTTCTTGGAATTTTCCCGATGTTTCATTAAGAT 184

QY 40 TLeGlyLysLysArgArgAlaThrArgPheThrPhePProArgArgGlyThrPro 58
DB 185 ACATGGCAAGAAAGAAAGAGCAACAGATTCTGGATCCAGAGGAGCAACCA 239

RESULT 28

LOCUS BF679165 629 bp mRNA linear EST 21-DEC-2000
DEFINITION 602153737P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294761 5', mRNA sequence.

ACCESSION

BF679165

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 629)

AUTHORS

NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE

Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgarbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
CNA Library Preparation: CLONETECH Laboratories, Inc.

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Plate: L1CM1144 row: j column: 10
High quality sequence stop: 627.
Location/Qualifiers

FEATURES

source

```
1. 629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294761"
/lab_host="DH10B (TI phage-resistant)"
/clone_id="NIH_MGC_83"
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech); Site:1: Still (ggccgcctgcgc); Site:2: Still (ggccatcagcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
```

ORIGIN

Alignment Scores:

Pred. No.:	2.96e-25	Length:	629
Score:	268.00	Matches:	55
Percent Similarity:	94.83%	Conservative:	0
Best Local Similarity:	94.83%	Mismatches:	3
Query Match:	87.30%	Indels:	1
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679165 (1-629)

QY 1 MetGlnMetPheProPheSerProLeuPhePhePheGlnLeuLeuGlnSerSer 20
DB 67 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCTTCAATGCTAAACAAAGCTCC 126

QY 21 ArgArgLeuGluHisThr-PheValPheLeuArganPheSerLeuMetLeuLeuArgTy 40
DB 127 AGAAGGCTGGAACATACCTTTGTTCTTCTTGGAATTTTCCCGATGTTTCATTAAGAT 186

QY 41 TLeGlyLysLysArgArgAlaThrArgPheThrPhePProArgArgGlyThrPro 58
DB 187 AT-GGCAAGAAAGAAAGAGCAACAGATTCTGGATCCAGAGGAGCAACCA 239

RESULT 29

LOCUS BF680840 803 bp mRNA linear EST 21-DEC-2000
DEFINITION 602156088P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4297197 5', mRNA sequence.

ACCESSION

BF680840

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 803)

AUTHORS

NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE

Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgarbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
CNA Library Preparation: CLONETECH Laboratories, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM150 row: 0 column: 22
High quality sequence stop: 689.
Location/Qualifiers

FEATURES

SOURCE

```
1. .803
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4297197"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgacctggcc); Site 2: SfiI
(ggcattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCGCATTTATGCGC-3'
and 3' adaptor sequence:
5'-ATCTAGAGCGCGGCGCGGCGGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."
```

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5.38e-25	267.00	803	57	0	1	2	0
Percent Similarity:	96.61%						
Best Local Similarity:	96.61%						
Query Match:	86.97%						

US-10-031-158-14 (1-58) x BF680840 (1-803)

Qy 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLyGlnSerSer 20
Db 67 ATGAGAGTGTTCCTCCCAAGCCCACTATTTCTTCCTCAATG-CTGAAACAAAGCTCC 125

Qy 21 ArgArgLeuGlnIshTrpPheValPheLeuArgAsnPheSerLeuMetLeuArgTy 40
Db 126 AGAAGGCTGGAAACATCTTGTCTTCTTGAGAAATTTTCCCTGATCTTATTAAAGATAC 185

Qy 41 IlleglYlyeLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 186 ATAGGCAAAAGAAAGAGCAACAGATTCTGGATCCCAAGAGGGAGAACACCA 240

RESULT 30

BI836020 772 bp mRNA linear EST 04-OCT-2001

LOCUS 603084004F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5223193 5',
mRNA sequence.

ACCESSION BI836020
VERSION BI836020.1 GI:15947570
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 772)
NIH-MGC http://img.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM1561 row: 9 column: 02

High quality sequence stop: 768.
Location/Qualifiers

FEATURES

SOURCE

```
1. .772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5223193"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/notes="Organ: pooled pancreas and spleen; Vector:
pCW-SORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dt primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2.39e-24	262.00	772	54	2	2	2	0
Percent Similarity:	93.33%						
Best Local Similarity:	90.00%						
Query Match:	85.34%						

US-10-031-158-14 (1-58) x BI836020 (1-772)

Qy 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLyGlnSerSer 20
Db 427 ATGAGAGTGTTCCTCCCAAGCCCACTATTTCTTCCTCAATGCTGAAACAAAGCTCC 486

Qy 21 ArgArgLeuGlnIshTrpPheValPheLeuArgAsnPheSerLeuMetLeuArgTy 40
Db 487 AGAAGGCTGGAAACATCTTGTCTTCTTGAGAAATTTTCCCAATATTATTAAAGATA 546

Qy 40 rIlleglYlyeLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
Db 547 CATTGGCAACAAAGAAAGAGCAACAGATTCTGGATCCCAAGAGGGAGAACACCA 602

RESULT 31

BF673289 808 bp mRNA linear EST 21-DEC-2000

LOCUS 602136139F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4272944 5',
mRNA sequence.

ACCESSION BF673289
VERSION BF673289.1 GI:11947184
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 808)
NIH-MGC http://img.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LCM1087 row: m column: 09

FEATURES

SOURCE

```
1. .808
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"  
/clone="IMAGE:4272944"  
/lab_host="DH10B (T1 phage-resistant)"  
/clone_lib="NIH_MGC_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site_1: SfiI (ggcgctcgcc); Site_2: SfiI  
(ggcattatggcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."  
  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.33e-24 Length: 808  
Score: 259.00 Matches: 56  
Percent Similarity: 93.44% Conservative: 1  
Best Local Similarity: 91.80% Mismatches: 1  
Query Match: 84.36% Indels: 3  
DB: 2 Gaps: 0  
  
US-10-031-158-14 (1-58) x BF673289 (1-808)  
  
Qy 1 MetGlnMetPheProPseProLeuPhePheLeuGln-LeuLeuLyGlnSerSe 20  
Db 286 ATGCGAGTGTTCCTCCCAAGCCACTATTTCTTCTTCAATGCTGGAACAAAGCTC 345  
Qy 20 TARGTgLeuGlnIsthPhe-ValPheLeuArgAnPheSerLeuMetLeuArgT 40  
Db 346 CAGAGGCTGGAAACATCTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGAT 405  
Qy 40 Yrlle-GlyLyLyArgArgAlaThrArgPheTrpAppProArgArgGlyThrPro 58  
Db 406 ACATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAAACACCA 462  
  
RESULT 32  
BF674457 857 bp mRNA linear EST 21-DEC-2000  
LOCUS 602137231P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4273734 5',  
DEFINITION mRNA sequence.  
ACCESSION BF674457  
VERSION BF674457.1 GI:11948352  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS 1 (bases 1 to 857)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Email: cga@bbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L10M1089 row: n column: 07  
High quality sequence stop: 638.  
Location/Qualifiers  
1..857  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:4273734"  
/lab_host="DH10B (T1 phage-resistant)"  
/lab_host="DH10B (T1 phage-resistant)"
```

```
/clone_lib="NIH_MGC_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site_1: SfiI (ggcgctcgcc); Site_2: SfiI  
(ggcattatggcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."  
  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.29e-23 Length: 857  
Score: 253.00 Matches: 56  
Percent Similarity: 96.55% Conservative: 0  
Best Local Similarity: 96.55% Mismatches: 2  
Query Match: 82.41% Indels: 2  
DB: 2 Gaps: 0  
  
US-10-031-158-14 (1-58) x BF674457 (1-857)  
  
Qy 1 MetGlnMetPheProPseProLeuPhePheLeuGlnLeuLeuLyGlnSerSe 20  
Db 63 ATGCGAGTGTTCCTCCCAAGCCACTATTTCTTCTTCAATG-CTGAAACAAAGCTCC 121  
Qy 21 ARGArgLeuGlnIsthPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40  
Db 122 AGAAGGCTGGAAACATCTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATC 181  
Qy 41 lIeglyLyLyArgArgAlaThrArgPheTrpAppProArgArgGlyThrPro 58  
Db 182 AT-GCGAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAAACACCA 234  
  
RESULT 33  
BF679284 849 bp mRNA linear EST 21-DEC-2000  
LOCUS 602153477P1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294386 5',  
DEFINITION mRNA sequence.  
ACCESSION BF679284  
VERSION BF679284.1 GI:11953179  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS 1 (bases 1 to 849)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Email: cga@bbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L10M1143 row: j column: 19  
High quality sequence stop: 587.  
Location/Qualifiers  
1..849  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:4294386"  
/lab_host="DH10B (T1 phage-resistant)"  
/clone_lib="NIH_MGC_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site_1: SfiI (ggcgctcgcc); Site_2: SfiI
```

(ggccatcatgccc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Alignment Scores:

Pred. No.:	7,84e-23	Length:	849
Score:	251.00	Matches:	56
Percent Similarity:	96.55%	Conservative:	0
Best Local Similarity:	96.55%	Mismatches:	2
Query Match:	81.76%	Indels:	2
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679284 (1-849)

QY 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20

Db 65 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTTCATTCGTAACAAAGCTCC 124

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

Db 125 AGAAGGCTGGAAACATACCTT-GTCTCTTGAGAAATTTTCCCGAAGTATTAAAGTAC 183

QY 41 ILGilylylylyyArgArgAlaThrArgPheTrpAspProArgArgTyrPro 58

Db 184 AT-GCGAAGAAAGAAAGACACACGATTCGGATCCGAGAGGGAACCA 236

QY 184 AT-GCGAAGAAAGAAAGACACACGATTCGGATCCGAGAGGGAACCA 236

RESULT 34

LOCUS

CO647648 995 bp mRNA linear EST 23-JUN-2004

DEFINITION ILLUMIGEN MQC 40761 Katze MMPB2 Macaca mulatta cDNA clone

TR00 (Hs.385086), mRNA sequence.

ACCESSION

CO647648

VERSION

CO647648.1

KEYWORDS

EST.

SOURCE

Macaca mulatta (rhesus monkey)

ORGANISM

Macaca mulatta

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L.

FEATURES

Source

1. .995

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

ORIGIN

Pred. No.:	2,38e-22	Length:	995
Score:	248.00	Matches:	49
Percent Similarity:	89.66%	Conservative:	3
Best Local Similarity:	84.48%	Mismatches:	6
Query Match:	80.78%	Indels:	0
DB:	7	Gaps:	0

US-10-031-158-14 (1-58) x CO647648 (1-995)

QY 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20

Db 468 ATCCAGATGTTTACCCCAAGCCCACTATTTCTTCTTCATTCGTAACAAAGCTCC 527

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

Db 528 ACAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCGAAGTATTAGATAC 587

QY 41 ILGilylylylyyArgArgAlaThrArgPheTrpAspProArgArgTyrPro 58

Db 588 ATTGCGAAGAAAGAAAGACACACGATTCGTAATCCCAAGAGGGAACCA 641

RESULT 35

LOCUS

CB553506 663 bp mRNA linear EST 01-JUN-2003

DEFINITION MMSPP0044.C12 MMSPP Macaca mulatta cDNA, mRNA sequence.

CB553506

VERSION CB553506.1

KEYWORDS EST.

SOURCE Macaca mulatta (rhesus monkey)

ORGANISM Macaca mulatta

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopithecoidea; Macaca.

1 (bases 1 to 663)

Katze, M.G., Bunzger, R., Korth, M., Feldman, R., Amjadi, M. and

Holzman, T.

Expressed sequence tags from Rhesus macaque spleen

Unpublished (2002)

Contact: Holzman T

Katze Lab

University of Washington

Box 358070, Seattle, WA 98195-8070, USA

Tel: 206 732 6156

Fax: 206 732 6055

Email: ted@locke.ws.washington.edu

Similar to Genbank entry HSTRG10.X06775 Human aberrant mRNA from

TRG gamma gene with unarranged J(J)2 spliced onto C(G)2. 3/1991

Plate: MMSPP0044 row: C column: 12.

FEATURES

Location/Qualifiers

1. .663

/organism="Macaca mulatta"

/mol_type="mRNA"

/db_xref="taxon:9544"

/sex="male"

/cell_type="monocytes"

/dev_stage="adult"

/clone_lib="MMSPP"

/note="Organ: spleen"

ORIGIN

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 514 ATAGGCTGGAACTACCTTTGCTTCTTGAGAAATTTTCCAGATGTTATTATGAGATAC 573
Qy 41 -11eg1LybLybArgArgAlaThrArgPheTrpAspProArgArgLyThrPro 58
Db 574 CATGGCCAGAAAAGAAACAGACACAGTTCTGAATCCACAGAGGGGACACCA 628

RESULT 38
BF677916 670 bp mRNA linear EST 21-DEC-2000
LOCUS 60208473BF1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4248993 5',
DEFINITION mRNA sequence.
ACCESSION BF677916 GI:11951811
VERSION BF677916.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 670)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM1068 row: 9 column: 10
High quality sequence stop: 668.

FEATURES
source location/Qualifiers
1..670
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4248993"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN
Alignment Scores:
Pred. No.: 2,37e-19 Length: 670
Score: 224.00 Matches: 55
Percent Similarity: 94.83% Conservative: 0
Best Local Similarity: 94.83% Mismatches: 3
Query Match: 72.96% Indels: 3
DB: Gaps: 0
US-10-031-158-14 (1-58) x BF677916 (1-670)

Qy 1 MetGlnMetPheProPserProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
Db 65 ATGAGAGTGTTCCTCCCAAGCCCACTATTTTCTTCCTTCATG-CTGAAACAAAGCTCC 123
Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

Db 124 AGAAGCTGGAACTACCTT-GTCTTCTTGAGAAATTTTCCGATGTTATTAGATAC 182
Qy 41 11eg1LybLybArgArgAlaThrArgPheTrpAspProArgArgLyThrPro 58
Db 183 AT-GGCAGAAAAGAAAGACACACATTTGATGATCCAGAGGGGACACCA 235

RESULT 39
BG535408 827 bp mRNA linear EST 03-APR-2001
LOCUS 60256306BF1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4687960 5',
DEFINITION mRNA sequence.
ACCESSION BG535408
VERSION BG535408.1 GI:13526953
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 827)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM1501 row: 1 column: 17
High quality sequence stop: 609.

FEATURES
source location/Qualifiers
1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4687960"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 7,64e-19 Length: 827
Score: 221.00 Matches: 50
Percent Similarity: 88.71% Conservative: 5
Best Local Similarity: 80.65% Mismatches: 3
Query Match: 71.99% Indels: 4
DB: Gaps: 0
US-10-031-158-14 (1-58) x BG535408 (1-827)

Qy 1 MetGlnMetPheProPserProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
Db 489 ATGAGAGTGTTCCTCCCAAGCCCACTATTTTCTTCCTTCGATGCTGAAACAAACTCC 548
Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 549 AGAAGCTGGAACTACCTTTGCTTCTTGAGAAATTTTCCAGATGTTATTATGACAT 608
Qy 40 11eg1LybLybArgArgAlaThrArgPheTrpAspProArgArgLyThrPro 58

Db 609 ACATGGCGAAGAAACGACGACGACGATTCGGATCCGAGAGGGAGAACCA 666

RESULT 40
BF370026 572 bp mRNA linear EST 24-NOV-2000
MR3-FN0004-120600-001-g12 FN0004 Homo sapiens CDNA, mRNA sequence.

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BF370026
MR3-FN0004-120600-001-g12 FN0004 Homo sapiens CDNA, mRNA sequence.
BF370026.1 GI:11331960
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 572)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL
MEDLINE
PubMed
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR3&t2=MR3-FN0004-
120600-001-g12&t3=2000-06-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 46.
Location/Qualifiers

FEATURES
source
1..572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="FN0004"
/note="Torgan: prostate normal; Vector: puc18; Site_1:
Sma1; Site_2: Sma1; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 1,97e-17 Length: 572
Score: 209.00 Matches: 44
Percent Similarity: 80.00% Conservative: 4
Best Local Similarity: 73.33% Mismatches: 10
Query Match: 68.08% Indels: 2
Gaps: 0

US-10-031-158-14 (1-58) x BF370026 (1-572)

OY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
|||||
:::|||||
7 ATGCAGATGTACCCCAAGCCACCTATATATCTTCCTCAGTGCGTGAACAAAGCTCC 66
|||||
21 ArgAglGluGlnHisThrPheValPheLeu--ArgAsnPheSerLeuMetLeuAArgT 40
|||||
67 AGAGGGCTGGACATTAACCTTGTGCTCTGTGAGATATTTCCTCGAATTAAGAT 126

OY 40 YrllGlyLylylGArgArgAlaThrArgPheTrpAppProAArgGlyThrPro 58
|||||
:::|||||
Db 127 ACATGGCGAAGAAAGAGAGCAACAGATTCGGATCCGAGAGGGAGAACCA 182

Search completed: December 8, 2004, 09:19:12
Job time: 1524 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2004, 07:51:31 ; Search time 56 Seconds

(without alignments)
736.174 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307
Sequence: 1 MOMPSPPLFPLQLKQSS.....RYIGKKRRATREMPDRGTP 58

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 353394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.epool/US10031158/runat_06122004_082700_15922/app_query.fasta_1.199
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pct -NOMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158 @CGN1_1.54@runat_06122004_082700_15922 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	426	4	US-09-513-999C-679
2	307	100.0	825	1	US-08-256-964A-18
3	279	90.9	1080	6	5260223-3
4	251	81.8	460	4	US-09-513-999C-8197
5	65.5	21.3	1789	4	US-09-142-108C-1
6	64	20.8	503	4	US-09-250-609-85
7	64	20.8	503	4	US-09-250-611-85
8	63.5	20.7	74962	4	US-09-685-853A-3
9	63	20.5	195	4	US-09-248-796A-10951
10	62.5	20.4	234	4	US-09-248-796A-11408
11	62.5	20.4	4070	3	US-09-302-812-1
12	62.5	20.4	4070	3	US-09-511-477-1

13	62.5	20.4	4070	3	US-09-511-507-1	Sequence 1, Appl1
14	61.5	20.0	2755	3	US-09-025-769B-274	Sequence 274, App
15	61.5	20.0	2755	4	US-09-490-070A-274	Sequence 274, App
16	61.5	20.0	2755	4	US-09-490-153-274	Sequence 274, App
17	61	19.9	339	4	US-09-250-609-79	Sequence 79, Appl
18	61	19.9	339	4	US-09-250-611-79	Sequence 79, Appl
19	61	19.9	399	4	US-09-250-609-91	Sequence 91, Appl
20	61	19.9	399	4	US-09-250-611-91	Sequence 91, Appl
21	61	19.9	420	4	US-09-250-609-89	Sequence 89, Appl
22	61	19.9	420	4	US-09-250-611-89	Sequence 89, Appl
23	61	19.9	424	4	US-09-250-609-77	Sequence 77, Appl
24	61	19.9	424	4	US-09-250-611-77	Sequence 77, Appl
25	61	19.9	424	4	US-09-250-609-90	Sequence 90, Appl
26	61	19.9	471	4	US-09-250-611-90	Sequence 90, Appl
27	61	19.9	540	4	US-09-250-609-95	Sequence 95, Appl
28	61	19.9	540	4	US-09-250-611-95	Sequence 95, Appl
29	61	19.9	836	4	US-09-495-050A-103	Sequence 103, App
30	61	19.9	2136	4	US-09-587-184-1	Sequence 1, Appl1
31	61	19.9	3798	4	US-09-688-188B-9	Sequence 9, Appl1
32	61	19.9	3798	4	US-09-291-417D-9	Sequence 9, Appl1
33	61	19.9	4266	3	US-09-651-011A-3	Sequence 3, Appl1
34	61	19.9	7295	2	US-08-487-826B-15	Sequence 15, Appl1
35	60.5	19.7	219	4	US-09-328-352-324	Sequence 324, Ap
36	60.5	19.7	504	4	US-09-134-000C-3042	Sequence 3042, Ap
37	60.5	19.7	1463	3	US-08-943-731-7	Sequence 7, Appl1
38	60.5	19.7	18609	3	US-08-943-731-1	Sequence 1, Appl1
39	60	19.5	2715	4	US-09-543-681A-192	Sequence 192, App
40	60	19.5	9840	3	US-09-534-638-1	Sequence 1, Appl1
41	60	19.5	30350	4	US-10-118-328-3	Sequence 3, Appl1
42	59.5	19.4	357	4	US-09-513-999C-15620	Sequence 15620, A
43	59.5	19.4	1144	4	US-09-149-476-128	Sequence 128, App
44	59.5	19.4	1482	4	US-09-023-655-191	Sequence 191, App
45	59.5	19.4	2095	4	US-09-608-917A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-513-999C-679
; Sequence 679, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 679
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 267..425
US-09-513-999C-679
Alignment Scores: 1.64e-36
Pred. No.: 307.00
Score: 100.00%
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0
US-10-031-158-14 (1-58) x US-09-513-999C-679 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
 DB 94 ATGCAGATGTTTCCCAAGCCCACTATTCTTCCTCAATGCTGAACAAAGCTCC 153
 QY 21 ArgArgLeuGlnHisThrPheVal1PheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 154 AGAAGGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 213
 QY 41 IlegLYLyeLYsArgArgAlaThrArgPheTTPAspProArgGlyThrPro 58
 DB 214 ATTGGCAAGAAAGAGAGACACGATTTCTGGATCCCAAGAGGGAACACCA 267

RESULT 2

US-08-256-964A-18
 ; Sequence 18, Application US/08256964A
 ; Patent No. 5723309

GENERAL INFORMATION:

APPLICANT: BONNEVILLE, Marc
 TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T
 TITLE OF INVENTION: RECEPTORS BY CO-TRANSFECTION, USES OF THE PRODUCTS THUS
 TITLE OF INVENTION: OBTAINED
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Young & Thompson
 STREET: 745 South 23rd Street
 CITY: Arlington
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/256,964A
 FILING DATE: 14-SEP-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 92 14203
 FILING DATE: 25-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: PATCH, Andrew J.
 REGISTRATION NUMBER: Reg. No. 5723309 32,925
 REFERENCE/DOCKET NUMBER: BE 94/449
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703/521-2297
 TELEFAX: 703/685-0573
 TELEX: 248425

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
 LENGTH: 825 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:

NAME/KEY: CDS
 LOCATION: 1..825

PUBLICATION INFORMATION:
 DOCUMENT NUMBER: WO 94/12648

FILING DATE: 25-NOV-1993
 PUBLICATION DATE: 09-JUN-1994

US-08-256-964A-18

Alignment Scores:

Pred. No.: 4.18e-36 Length: 825
 Score: 307.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-10-031-158-14 (1-58) x US-08-256-964A-18 (1-825)
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
 DB 440 ATGCAGATGTTTCCCAAGCCCACTATTCTTCCTCAATGCTGAACAAAGCTCC 499
 QY 21 ArgArgLeuGlnHisThrPheVal1PheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
 DB 500 AGAAGGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 559
 QY 41 IlegLYLyeLYsArgArgAlaThrArgPheTTPAspProArgGlyThrPro 58
 DB 560 ATTGGCAAGAAAGAGAGACACGATTTCTGGATCCCAAGAGGGAACACCA 613

RESULT 3

5260223-3
 ; Patent No. 5260223

APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN,
 JOHN G., JR., STEPHEN H.; KRANGEL, MICHAEL S.
 TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA

T CELL RECEPTOR

NUMBER OF SEQUENCES: 4
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/187,698

FILING DATE: 29-APR-1988

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 115,256

FILING DATE: 29-OCT-1987

APPLICATION NUMBER: 16,252

FILING DATE: 19-FEB-1987

APPLICATION NUMBER: 882,100

FILING DATE: 03-JUL-1986

SEQ ID NO:3:

LENGTH: 1080

5260223-3

Alignment Scores:

Pred. No.: 8.8e-32 Length: 1080
 Score: 279.00 Matches: 53
 Percent Similarity: 96.55% Conservative: 3
 Best Local Similarity: 91.38% Mismatches: 2
 Query Match: 90.88% Indels: 0
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x 5260223-3 (1-1080)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 452 ATGCAGATGTTTCCCAAGCCCACTATTCTTCCTCAATGCTGAACAAAGCTCC 511

QY 21 ArgArgLeuGlnHisThrPheVal1PheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

DB 512 AGAAGGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCAATATTATTAAAGATAC 571

QY 41 IlegLYLyeLYsArgArgAlaThrArgPheTTPAspProArgGlyThrPro 58

DB 572 ATTGGCAAGAAAGAGAGACACGATTTCTGGATCCCAAGAGGGAACACCA 625

RESULT 4

US-09-513-999C-8197

; Sequence 8197, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

APPLICANT: Dumas, Milne Edwards, J.B.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 FILE REFERENCE: 59,US2,REG
 CURRENT APPLICATION NUMBER: US/09/513,999C
 CURRENT FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/122,487
 PRIOR FILING DATE: 1999-02-26

```

; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8197
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 327
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 345
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 421
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n=a, g, c or t
; US-09-513-999C-8197

```

```

Alignment Scores:
Pred. No.: 3,78e-28      Length: 460
Score: 251.00           Matches: 56
Percent Similarity: 96.53%      Conservative: 0
Best Local Similarity: 96.55%      Mismatches: 2
Query Match: 81.76%             Indels: 2
DB: 4                      Gaps: 0

```

US-10-031-158-14 (1-58) x US-09-513-999C-8197 (1-460)

```

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInserSer 20
    |||||
DB 200 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTCAATGCTGAACAAAGCTC- 258
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTy 40
    |||||
DB 259 ACAAGGCTGGAACATCTCTTGCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 318
QY 41 IleglylveLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
    |||||
DB 319 ATTGGGAARAAAGAGAGCAACAGTC-TGGGATCCAGAGGAGGAAACCA 371

```

RESULT 5

US-09-142-108C-1/c
; Sequence 1, Application US/09142108C

```

; Patent No. 6774285
; GENERAL INFORMATION:
; APPLICANT: Bruggliera, Filipa
; APPLICANT: Holton, Timothy A.
; APPLICANT: Michael, Michael Z.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID PATHWAY ENZYMES
; FILE REFERENCE: 11658
; CURRENT APPLICATION NUMBER: US/09/142,108C
; CURRENT FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: PN8386
; PRIOR FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1789
; TYPE: DNA
; ORGANISM: Petunia x hybrida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50) .. (1588)
; US-09-142-108C-1

```

Alignment Scores:

```

Pred. No.: 9.03      Length: 1789
Score: 65.50         Matches: 18
Percent Similarity: 63.04%      Conservative: 11
Best Local Similarity: 39.13%      Mismatches: 14
Query Match: 21.34%             Indels: 4
DB: 4                      Gaps: 1

```

US-10-031-158-14 (1-58) x US-09-142-108C-1 (1-1789)

```

QY 4 PheProProSerProLeuPhePheLeuGlnLeuLeuYsgInserSerArgArgLeu 23
    |||||
DB 912 TTCCTTCATCAT-ATCCGATCATCTTTTAAAGAGATCAAGATCAACAAATC 854
QY 24 GlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyrlleglylve 43
    |||||
DB 853 TTTCAAT-----TTTCCAAAAATTTTACCCTTAATGCTCTTCAAGTATATCACTCA 803
QY 44 LysArgArgAlaThrArg 49
    |||||
DB 802 GAAGCGCTCGAAACGCC 785

```

RESULT 6

US-09-250-609-85
; Sequence 85, Application US/09250609A

```

; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-250-609-85

```

Alignment Scores:

```

Pred. No.: 2.5      Length: 503
Score: 64.00       Matches: 13
Percent Similarity: 59.26%      Conservative: 3
Best Local Similarity: 48.15%      Mismatches: 11
Query Match: 20.85%             Indels: 0
DB: 4                      Gaps: 0

```

US-10-031-158-14 (1-58) x US-09-250-609-85 (1-503)

```

QY 31 ArgAsnPheSerLeuMetLeuLeuArgTyrlleglylveLysArgArgAlaThrArgPhe 50
    |||||
DB 246 AGGAGCTTCGGGCTAGCTGCTAAGTGAAGAAATTTGTCACTCTGCCCAAGGTGC 305
QY 51 TyrPheProArgArgGlyThr 57
    |||||
DB 306 TCGACGCCAAGAGAGGCACT 326

```

RESULT 7

US-09-250-611-85
; Sequence 85, Application US/09250611

```

; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Bassett, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 503
; TYPE: DNA

```

```

US-09-250-611-85
; ORGANISM: Homo sapiens
US-09-250-611-85

Alignment Scores:
Pred. No.: 2.5 Length: 503
Score: 64.00 Matches: 13
Percent Similarity: 59.26% Conservative: 3
Best Local Similarity: 48.15% Mismatches: 11
Query Match: 20.85% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-611-85 (1-503)

QY 31 ArgAnpHeSerLeuMetLeuLeuArgTyrIleGlyLysIbYbArgAlaThrArgPhe 50
Db 246 AGGACCTTCGGGCTCGAGCTTCTTAAGTGTGAAGAAGAAATTGCATCTCGCCAGGTGC 305
QY 51 TTPApPProArgArgIYThr 57
Db 306 TGGCAGCCAAAGAGGCACT 326

RESULT 8
US-09-685-853A-3
; Sequence 3, Application US/09685853A
; Patent No. 6479270
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS.
; FILE REFERENCE: CLO00871
; CURRENT APPLICATION NUMBER: US/09/685, 853A
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/182,194
; PRIOR FILING DATE: 2000-02-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 74962
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(74962)
; OTHER INFORMATION: n = A,T,C or G
US-09-685-853A-3

Alignment Scores:
Pred. No.: 3.55e+03 Length: 74962
Score: 63.50 Matches: 22
Percent Similarity: 47.62% Conservative: 8
Best Local Similarity: 34.92% Mismatches: 15
Query Match: 20.68% Indels: 18
DB: 4 Gaps: 3

US-10-031-158-14 (1-58) x US-09-685-853A-3 (1-74962)

QY 10 PhePhePheLeuGLeuLeuLysGlnSerArg-----Arg 22
Db 44953 TTTTITTTTTCCCTTTACTTAAGATGAGTTCACAAACAGTAGATGTGTGACTCTCGTTC 45012
QY 23 LeuGluH1srThrPhe-ValPheLeuArgAnpHeSer-----Le 35
Db 45013 ATAGACCAATCACTTCGTTTTATATGTTGCTCTTAGACTTTGACCTCTCTGAAA 45072
QY 35 uMetLeuLeuArgTyrIleGlyLysIbYbArgArgAla-----ThrArgPheTrpAs 52
Db 45073 TGCTCTGTAGTTCATGACGAGAGTCAACAGACCACTTAGAGAACAATCTTCTGTGC 45132
QY 52 pProArg 54
Db 45133 TTAGAGA 45139

```

```

US-09-248-796A-10951
/ Sequence 10951, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 10951
/ LENGTH: 195
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-10951

Alignment Scores:
Pred. No.: 0.92 Length: 195
Score: 63.00 Matches: 10
Percent Similarity: 92.86% Conservative: 3
Best Local Similarity: 71.43% Mismatches: 1
Query Match: 20.52% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-248-796A-10951 (1-195)

QY 3 MetPheProPserProLeuPhePhePheLeuGlnIleu 16
Db 131 TTATTCCCTCCACCCCTTTT TTTT TTTTATTCACCTTTTA 172

RESULT 10
US-09-248-796A-11408
/ Sequence 11408, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 11408
/ LENGTH: 234
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-11408

Alignment Scores:
Pred. No.: 1.41 Length: 234
Score: 62.50 Matches: 17
Percent Similarity: 56.82% Conservative: 8
Best Local Similarity: 38.64% Mismatches: 18
Query Match: 20.36% Indels: 1
DB: 4 Gaps: 1

US-10-031-158-14 (1-58) x US-09-248-796A-11408 (1-234)

QY 9 LeuPhePhePheLeuGlnIleuLeuIysGlnSerSerArgLeuGlnHisThr---Phe 27
Db 31 TTAATTAATCTTCATTTCTGATTAAGAAAGAGGAGAAATTTAGAGAACTCAATTTT 90

28 ValPheLeuAlaGlnPheSerLeuMetLeuLeuArgTyrIleGlyIleValArgArgAla 47

```

```

Db      91 TTTTITTTTTAGATTCATTCTTCTTTGTTGTATTGAAGAAAAAACCATA 150
Qy      48 ThrArgPheTp 51
           |||||
           :~::~:~::
Db      151 CCAAGATTTT 162

RESULT 11
US-09-302-812-1
; Sequence 1, Application US/09302812B
; Patent No. 6333148
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOPHYROLASE (PARC) ENZYME COMPOSITIONS AND ANTIBODIES IMMUNOREACTIVE TO THEM
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE TO THEM
; FILE REFERENCE: NIND 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; EARLIER FILING DATE: 1999-04-30
; PRIOR FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 1
; LENGTH: 4070
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
US-09-302-812-1

Alignment Scores:
Pred. No.:          Length:        4070
Score:              Matches:         17
Percent Similarity: 62.50%
Best Local Similarity: 53.85%
Query Match:       Conservative:    4
                  Mismatches:     17
                  Indels:          1
                  Gaps:            1
                  DB:               1
                        3

US-10-031-158-14 (1-58) x US-09-302-812-1 (1-4070)

Qy      1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
           |||||||
Dd      3393 ATGGAGTGCCCTTCGCCCATGCAGCTGTCAGTAGCATCGCCCTTCATCCGACTTA 3352
           |||||
Qy      21 ArgArgLeuGluHisThr---PheValPheLeuArgAsnPheserIleuMetLeuLeu 38
           |||||
Dd      3353 CATGAGTGCAGACACTTTGTTCCTTTTTCATTTCATTCAGCCCTGATTTCTTTTA 3409
           |||||

RESULT 12
US-09-511-477-1
; Sequence 1, Application US/09511477
; Patent No. 6337202
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOPHYROLASE (PARC) ENZYME COMPOSITIONS AND ANTIBODIES IMMUNOREACTIVE TO THEM
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE TO THEM
; FILE REFERENCE: NIND 201
; CURRENT APPLICATION NUMBER: US/09/511,477
; EARLIER FILING DATE: 2000-02-23
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 1
; LENGTH: 4070
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:

```

```

US-09-511-477-1
Alignment Scores:
Pred. No.: 80.7 Length: 4070
Score: 62.50 Matches: 17
Percent Similarity: 53.85% Conservative: 4
Best Local Similarity: 43.59% Mismatches: 17
Query Match: 20.36% Indels: 1
DB: 3 Gaps: 1

US-10-031-158-14 (1-58) x US-09-511-477-1 (1-4070)
QY 1 MetGlnMetPheProPserProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
Db 3293 ATGCAAGTCCTTCGCGCCATGCAGCTGTCAATCATGTGGCCCTCCATCCGACTTA 3352
QY 21 ArgArgLeuGlnHisThr---PheValPheLeuArgAsnPheSerLeuMetLeu 38
Db 3353 CATGAGCTGCAGACATCTTTGTTCTTTTTCATTTCATTCCAGCCGATCTTTTA 3409

RESULT 13
US-09-511-507-1
Sequence 1, Application US/09511507
Patent No. 6395543
GENERAL INFORMATION:
APPLICANT: JACOBSON, Myron K.
APPLICANT: JACOBSON, Elaine L.
APPLICANT: AME, Jean-Christophe
APPLICANT: LIN, Winston
TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) EN
TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
FILE REFERENCE: NIAD 201
CURRENT APPLICATION NUMBER: US/09/511,507
CURRENT FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/302,812
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 1
LENGTH: 4070
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
US-09-511-507-1

Alignment Scores:
Pred. No.: 80.7 Length: 4070
Score: 62.50 Matches: 17
Percent Similarity: 53.85% Conservative: 4
Best Local Similarity: 43.59% Mismatches: 17
Query Match: 20.36% Indels: 1
DB: 3 Gaps: 1

US-10-031-158-14 (1-58) x US-09-511-507-1 (1-4070)
QY 1 MetGlnMetPheProPserProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
Db 3293 ATGCAAGTCCTTCGCGCCATGCAGCTGTCAATCATGTGGCCCTCCATCCGACTTA 3352
QY 21 ArgArgLeuGlnHisThr---PheValPheLeuArgAsnPheSerLeuMetLeu 38
Db 3353 CATGAGCTGCAGACATCTTTGTTCTTTTTCATTTCATTCCAGCCGATCTTTTA 3409

RESULT 14
US-09-025-7698-274/c
Sequence 274, Application US/090257698
Patent No. 6300064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Knappik, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon

```

```
/
/ APPLICANT: Plueckthun, Andreas
/ TITLE OF INVENTION: Protein/(Poly)peptide libraries
/ NUMBER OF SEQUENCES: 373
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: James F. Haley, Jr., Esq. c/o Fish & Neave
/ STREET: 1251 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10021
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/025,769B
/ FILING DATE: 18-FEB-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 95 11 3021.0
/ FILING DATE: 18-AUG-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: James F. Haley, Jr., Esq.
/ REGISTRATION NUMBER: 27,794
/ REFERENCE/DOCKET NUMBER: MORPHO/5
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212)596-9000
/ TELEFAX: (212)596-9090
/ INFORMATION FOR SEQ ID NO: 274:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2755 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: other nucleic acid
/ DESCRIPTION: /desc = "synthetic vector"
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..509
/ OTHER INFORMATION: /product= "gIIIp ss, myc tag, amber"
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: complement (1853..2509)
/ OTHER INFORMATION: /product= "cat resistance"
/
/ US-09-025-769B-274
/
/ Alignment Scores:
/ Pred. No.: 65.3 Length: 2755
/ Score: 61.50 Matches: 15
/ Percent Similarity: 52.83% Conservative: 13
/ Best Local Similarity: 28.30% Mismatches: 8
/ Query Match: 20.03% Indels: 17
/ DB: 3 Gaps: 3
/
/ US-10-031-158-14 (1-58) x US-09-025-769B-274 (1-2755)
/
/ QY 2 GlnMetPheProPseSerProLeuPhePhePheLeuInleuLysGlnSerSerArg 21
/ DB 602 CAGGCCGCGCCCCCCCCCTT-----TTAATTAAACGCGACACAAA 561
/
/ QY 22 ArgLeuGluHisThrPheValPheLeuArgAsnPheSer----- 34
/ DB 560 AAATATCGCACAT-----CTCGGCATTTCCTTCACCTTCACAGTCAAGCTTATCA 510
/
/ QY 35 ---LeuMetLeuLeuArgTyrIleGlyLysLysArgArg 46
/ DB 509 AGACTCCTATTACGACAGTATGTTAGCAAAAGTGAATA 471
/
/ RESULT 15
/ US-09-490-070A-274/C
/ Sequence 274, Application US/09490070A
/ Patent No. 6696248
```

```
/
/ GENERAL INFORMATION:
/ APPLICANT: Knappik, Achim
/ PACK, Peter
/ Ilag, Vic
/ Ge, Liming
/ Moroney, Simon
/ Plueckthun, Andreas
/ TITLE OF INVENTION: Protein/(Poly)peptide libraries
/ NUMBER OF SEQUENCES: 373
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Colin G. Sandercock, Esq. c/o Heller Ehrman
/ STREET: 1666 K Street, N.W., Suite 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20006
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/490,070A
/ FILING DATE: 24-Jan-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 95 11 3021.0
/ FILING DATE: 18-AUG-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Colin G. Sandercock, Esq.
/ REGISTRATION NUMBER: 31,298
/ REFERENCE/DOCKET NUMBER: 37629-0005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 912-2000
/ TELEFAX: (202) 912-2020
/ INFORMATION FOR SEQ ID NO: 274:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2755 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: other nucleic acid
/ DESCRIPTION: /desc = "synthetic vector"
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..509
/ OTHER INFORMATION: /product= "gIIIp ss, myc tag, amber"
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: complement (1853..2509)
/ OTHER INFORMATION: /product= "cat resistance"
/ SEQUENCE DESCRIPTION: SEQ ID NO: 274:
/
/ US-09-490-070A-274
/
/ Alignment Scores:
/ Pred. No.: 65.3 Length: 2755
/ Score: 61.50 Matches: 15
/ Percent Similarity: 52.83% Conservative: 13
/ Best Local Similarity: 28.30% Mismatches: 8
/ Query Match: 20.03% Indels: 17
/ DB: 4 Gaps: 3
/
/ US-10-031-158-14 (1-58) x US-09-490-070A-274 (1-2755)
/
/ QY 2 GlnMetPheProPseSerProLeuPhePhePheLeuInleuLysGlnSerSerArg 21
/ DB 602 CAGGCCGCGCCCCCCCCCTT-----TTAATTAAACGCGACACAAA 561
/
/ QY 22 ArgLeuGluHisThrPheValPheLeuArgAsnPheSer----- 34
/ DB 560 AAATATCGCACAT-----CTCGGCATTTCCTTCACCTTCACAGTCAAGCTTATCA 510
```



```

01 5 ---LeumleuLeuAaGyrrileGlylyblybAarg 46
02 |||:|||||:|||||:|||||:|||||:|||||:|
03 Db 509 AGACTCTATTACGCAATGTTAGCAACGTGAAAA 471
04
05 RESULT 16
06 US-09-490-153-274/c
07 Sequence 274, Application US/09490153
08 Patent No. 6706484
09 GENERAL INFORMATION:
10 APPLICANT: Knappik, Achim
11 Pack, Peter
12 Ilag, Vic
13 Ge, Liming
14 Moroney, Simon
15 Plueckthun, Andreas
16 TITLE OF INVENTION: Protein/(Poly)peptide libraries
17 NUMBER OF SEQUENCES: 373
18 CORRESPONDENCE ADDRESS:
19 ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
20 STREET: 1251 Avenue of the Americas
21 CITY: New York
22 STATE: New York
23 COUNTRY: USA
24 ZIP: 10021
25
26 COMPUTER READABLE FORM:
27 MEDIUM TYPE: Floppy disk
28 COMPUTER: IBM PC compatible
29 OPERATING SYSTEM: PC-DOS/MS-DOS
30 SOFTWARE: Patentin Release #1.0, Version #1.30 (ERO)
31 CURRENT APPLICATION DATA:
32 APPLICATION NUMBER: US/09/490,153
33 FILING DATE: 24-Jan-2000
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: US/09/025,769B
36 FILING DATE: 18-FEB-1998
37 APPLICATION NUMBER: EP 95 11 3021.0
38 FILING DATE: 18-AUG-1995
39 ATTORNEY/AGENT INFORMATION:
40 NAME: James F. Haley, Jr., Esq.
41 REGISTRATION NUMBER: 27,794
42 REFERENCE/DOCKET NUMBER: MORPHO/5
43 TELECOMMUNICATION INFORMATION:
44 TELEPHONE: (212)596-9000
45 TELEFAX: (212)596-9090
46 INFORMATION FOR SEQ ID NO: 274:
47 SEQUENCE CHARACTERISTICS:
48 LENGTH: 2755 base pairs
49 TYPE: nucleic acid
50 STRANDEDNESS: double
51 TOPOLOGY: circular
52 MOLECULE TYPE: other nucleic acid
53 DESCRIPTION: /desc = "synthetic vector"
54 FEATURE:
55 NAME/KEY: CDS
56 LOCATION: 3..509
57 OTHER INFORMATION: /product= "giip ss, myc tag, amber
58 codon"
59
60 FEATURE:
61 NAME/KEY: CDS
62 LOCATION: complement (1853..2509)
63 OTHER INFORMATION: /product= "cat resistance"
64 SEQUENCE DESCRIPTION: SEQ ID NO: 274:
65
66 US-09-490-153-274
67
68 Alignment Scores:
69 Pred. No.: 65.3 Length: 2755
70 Score: 61.50 Matches: 15
71 Percent Similarity: 52.83% Conservative: 13
72 Best Local Similarity: 20.03% Mismatches: 8
73 Query Match: 20.03% Indels: 17
74 Gaps: 3
75
76 US-10-031-158-14 (1-58) x US-09-490-153-274 (1-2755)

```

```

Oy      GlMeEPhaPrProSerProLeuNhePhePheLeuGlnLeuLysGlnSerSerArg 22
Db      602 CAGCGCGCGCCCCCCCCCTT-----TTAATTAAACGGCAGACAAA 561

Oy      22 ArgLeuGlnLisThrPheValPheLeuArgAnPheSer 34
Db      560 AAAATGTCCACAAT-----CTGGCCCATTTTTCCTTCAACAGGTCAAGTTATCA 510

Oy      35 ---LeuMetLeuLeuArgTyrTlleGlyLysLysArgArg 46
Db      509 AGACTCTTATTACGACGATGTGTGACAAAGTAGAAAA 471

RESULT 17
US-09-250-609-79
; Sequence 79, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-79

Alignment Scores:
Pred. No.:      3.99      Length:      339
Score:          61.00      Matches:      18
Percent Similarity: 48.94%      Conservative: 5
Best Local Similarity: 38.30%      Mismatches: 14
Query Match:     19.87%      Indels:      10
DB:              4          Gaps:      2

US-10-031-158-14 (1-58) x US-09-250-609-79 (1-339)

Oy      13 LeuGlnLeuLeuLys-----GlnSerSerArgArgLeuGlnLisThrPheValPheLeu 30
Db      197 CTCACGTTGTAGAAAGCCTGACAGAGCGGAGAGAAAGATT----- 235

Oy      31 ArgAsnPheSerLeuMetLeuLeuArgTyrTlleGlyLysLysArgArgAlaThrArgPhe 50
Db      236 ---AGCTTGCGGCTGAGCTTGCTTAAGTGGAAGAAAGAAATTGTCACTCTGGCCAGAGTGC 292

Oy      51 TrpAspProArgArgGlyThr 57
Db      293 TGGCAGCCAAAGAGAGGCACT 313

RESULT 18
US-09-250-611-79
; Sequence 79, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-611-79

Alignment Scores:

```

Pred. No.: 3,99 Length: 339
Score: 61.00 Matches: 18
Percent Similarity: 48.94% Conservative: 5
Best Local Similarity: 38.30% Mismatches: 14
Query Match: 19.87% Indels: 10
DB: 4 Gaps: 2

US-10-031-158-14 (1-58) x US-09-250-611-79 (1-339)

QY 13 LeuGlnLeuLeuLys-----GlnSerSerArgArgLeuGlnHisThrPheValPheLeu 30

DB 197 CTCGACCTTTAGAGGCTGACACAGGGGAGGAGAACTT----- 235

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 236 ---AGCTTGGGCTGAGCTTGCCTAAGCTGGAAGAAGAAATTGCACTCTGCCCGAGTGC 292

QY 51 TTPAspProArgArgGlyThr 57

DB 293 TGGCAGCCCAAGAGAGGCACT 313

RESULT 19

US-09-250-609-91
Sequence 91, Application US/09250609A
Patent No. 6458943

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210002
CURRENT APPLICATION NUMBER: US/09/250,609A

CURRENT FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 91

LENGTH: 399
TYPE: DNA

ORGANISM: Homo sapiens
US-09-250-609-91

Alignment Scores:

Pred. No.: 5,02 Length: 399
Score: 61.00 Matches: 12

Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11

Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-91 (1-399)

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 227 AAGAGCTTGGGCTGAGCTTGCCTAAGCTGGAAGAAGAAATTGCACTCTGCCCGAGTGC 286

QY 51 TTPAspProArgArgGlyThr 57

DB 287 TGGCAGCCCAAGAGAGGCACT 307

RESULT 20

US-09-250-611-91
Sequence 91, Application US/09250611
Patent No. 6528283

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210001
CURRENT APPLICATION NUMBER: US/09/250,611

CURRENT FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 91

LENGTH: 399
TYPE: DNA

ORGANISM: Homo sapiens
US-09-250-611-91

Alignment Scores:

Pred. No.: 5,02 Length: 399
Score: 61.00 Matches: 12

Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11

Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-611-91 (1-399)

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 227 AAGAGCTTGGGCTGAGCTTGCCTAAGCTGGAAGAAGAAATTGCACTCTGCCCGAGTGC 286

QY 51 TTPAspProArgArgGlyThr 57

DB 287 TGGCAGCCCAAGAGAGGCACT 307

RESULT 21

US-09-250-609-89
Sequence 89, Application US/09250609A
Patent No. 6458943

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210002
CURRENT APPLICATION NUMBER: US/09/250,609A

CURRENT FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 89

LENGTH: 420
TYPE: DNA

ORGANISM: Homo sapiens
US-09-250-609-89

Alignment Scores:

Pred. No.: 5,4 Length: 420
Score: 61.00 Matches: 12

Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11

Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-89 (1-420)

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 230 AAGAGCTTGGGCTGAGCTTGCCTAAGCTGGAAGAAGAAATTGCACTCTGCCCGAGTGC 289

QY 51 TTPAspProArgArgGlyThr 57

DB 290 TGGCAGCCCAAGAGAGGCACT 310

RESULT 22

US-09-250-611-89
Sequence 89, Application US/09250611
Patent No. 6528283

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210001
CURRENT APPLICATION NUMBER: US/09/250,611

CURRENT FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 89

LENGTH: 420
TYPE: DNA

```

; ORGANISM: Homo sapiens
US-09-250-611-89

Alignment Scores:
Pred. No.:          5.4           Length:      420
Score:              61.00         Matches:     12
Percent Similarity: 59.26%       Conserves:    4
Best Local Similarity: 44.44%   Mismatches:  11
Query Match:        19.87%      Indels:       0
DB:                  4           Gaps:          0

US-10-031-158-14 (1-58) x US-09-250-611-89 (1-420)

OY      31 ArgAnpHeSeLeumEtLeuAArgTyrIleGlyLyIsLysArGaAlarHArGpHe 50
      ::::||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      230 AAGACCTTCGGCGCTGACTGTCTAAAGTGGAAGAAGAAATTGTCACTTCGCCAGGTGC 289
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      51 TrPaSPProArGaRgGLyThr 57
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      290 TGGCAGCCAAAGAGAGGCACT 310
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 23
US-09-250-609-77
; Sequence 77, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-77

Alignment Scores:
Pred. No.:          5.47           Length:      424
Score:              61.00         Matches:     12
Percent Similarity: 59.26%       Conservative:  4
Best Local Similarity: 44.44%   Mismatches:   11
Query Match:        19.87%      Indels:       0
DB:                  4           Gaps:          0

US-10-031-158-14 (1-58) x US-09-250-609-77 (1-424)

OY      31 ArgAnpHeSeLeumEtLeuAArgTyrIleGlyLyIsLysArGaAlarHArGpHe 50
      ::::||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      239 AAGACCTTCGGCGCTGACTGTCTAAAGTGGAAGAAGAAATTGTCACTTCGCCAGGTGC 289
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      51 TrPaSPProArGaRgGLyThr 57
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      299 TGGCAGCCAAAGAGAGGCACT 319
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 24
US-09-250-611-77
; Sequence 77, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Basset, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
; LENGTH: 424
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-09-250-611-77

Alignment Scores:
Pred. No.:          5.47           Length:        424
Score:              61.00          Matches:         12
Percent Similarity: 59.26%         Conservative:    4
Best Local Similarity: 44.44%       Mismatches:      11
Query Match:        19.87%         Indels:            0
DB:                  4             Gaps:            0

US-10-031-158-14 (1-58) x US-09-250-611-77 (1-424)

QY      31 ArgAnPheSeLeuMeLenuRgrYrILeGlyVylSaArgAlarHArGrPe 50
       ::::||||| ||||| |::: |||||
Db      229 AAGACCTTCGGGCTTAGCCTGTGCTAAGTGAGAAGAAATTGTCACTTGC CGCCAGGTGC 298
               ||||| |::: ||||| |::: |||||

QY      51 TrPaEPProARgArGLyThr 57
       :||| ||||| |::: |||||
Db      299 TGGCAGCCCAAAGAGAGCACACT 319
               ||||| |::: ||||| |::: |||||

RESULT 25
US-09-250-609-90
; Sequence 90, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-90

Alignment Scores:
Pred. No.:          6.35           Length:        471
Score:              61.00          Matches:         12
Percent Similarity: 59.26%         Conservative:    4
Best Local Similarity: 44.44%       Mismatches:     11
Query Match:        19.87%         Indels:            0
DB:                  4             Gaps:            0

US-10-031-158-14 (1-58) x US-09-250-609-90 (1-471)

QY      31 ArgAnPheSeLeuMeLenuRgrYrILeGlyVylSaArgAlarHArGrPe 50
       ::::||||| ||||| |::: |||||
Db      248 AAGACCTTCGGGCTTAGCCTGTGCTAAGTGAGAAGAAATTGTCACTTGC CGCCAGGTGC 307
               ||||| |::: ||||| |::: |||||

QY      51 TrPaEPProARgArGLyThr 57
       :||| ||||| |::: |||||
Db      308 TGGCAGCCCAAAGAGAGCACACT 328
               ||||| |::: ||||| |::: |||||

RESULT 26
US-09-250-611-90
; Sequence 90, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICATION: Basaset, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 471
; TYPE: DNA
```

```

: ORGANISM: Homo sapiens
US-09-250-611-90

Alignment Scores:
Pred. No.: 6.35 Length: 471
Score: 61.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11
Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-611-90 (1-471)
OY 31 ArgAenPheSerLeuMetLeuAaGlyrIlleGlyLybLybArGArGAlaThrArgPhe 50
Db 248 AAGAGCTTGCGGCTGTGAGCTTGCTAAGGTGAAGAAATGTCACTTGCGCCAGGTGC 307
OY 51 TTPAaPProArGArgGlyThr 57
Db 308 TGGCAGCCAAAGAGAGGCACT 328

RESULT 27
US-09-250-609-95
Sequence 95, Application US/09250609A
Patent No. 6458943
GENERAL INFORMATION:
APPLICANT: Byrne, Jennifer A.
TITLE OF INVENTION: Members of the D52 Gene Family
FILE REFERENCE: 1383.0210002
CURRENT APPLICATION NUMBER: US/09/250,609A
CURRENT FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 95
LENGTH: 540
TYPE: DNA
ORGANISM: Homo sapiens
US-09-250-609-95

Alignment Scores:
Pred. No.: 7.71 Length: 540
Score: 61.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11
Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-95 (1-540)
OY 31 ArgAenPheSerLeuMetLeuAaGlyrIlleGlyLybLybArGArGAlaThrArgPhe 50
Db 96 AAGAGCTTGCGGCTGTGAGCTTGCTAAGGTGAAGAAATGTCACTTGCGCCAGGTGC 155
OY 51 TTPAaPProArGArgGlyThr 57
Db 156 TGGCAGCCAAAGAGAGGCACT 176

RESULT 28
US-09-250-611-95
Sequence 95, Application US/09250611
Patent No. 6528283
GENERAL INFORMATION:
APPLICANT: Byrne, Jennifer A.
TITLE OF INVENTION: Members of the D52 Gene Family
FILE REFERENCE: 1383.0210001
CURRENT APPLICATION NUMBER: US/09/250,611
CURRENT FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 95
LENGTH: 540
TYPE: DNA

```

[illegible]


```

Oy      23 uGlnH1eThrPheValPheLeuArg 31
Db      1194 CTTTCTCTCTCTGTTCTCTCTCTGC 1170

RESULT 34
US-08-487-826B-15/c
Sequence 15, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chltnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESSES:
ADDRESS: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487, 826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 7295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-15

Alignment Scores:
Pred. No.: 308 Length: 7295
Score: 61.00 Matches: 14
Percent Similarity: 50.00% Conservative: 4
Best Local Similarity: 38.89% Mismatches: 10
Query Match: 19.87% Indels: 8
DB: 2 Gaps: 1
US-10-031-15B-14 (1-58) x US-08-487-826B-15 (1-7295)

Oy      8 ProlaunPhePheLeuGlnLeuLeuLysGlnSerArgArgLeuGlnH1eThrPhe 27
Db      5473 CCACCTTTTCTTTTCTT-----GCTTCGCCGACAAATGTC 5438

Oy      28 ValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyrIleGlyLys 43
Db      5437 CTACTCTTCATCAGTTTCTTTTGTGGTCTTCATATCGACAAA 5390

RESULT 35
US-09-328-352-3224

```

```

Sequence 3224, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3224
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3224

Alignment Scores:
Pred. No.: 2.55 Length: 219
Score: 60.50 Matches: 11
Percent Similarity: 70.00% Conservative: 17
Best Local Similarity: 27.50% Mismatches: 7
Query Match: 19.71% Indels: 5
DB: 4 Gaps: 2

US-10-031-158-14 (1-58) x US-09-328-352-3224 (1-219)

Qy 13 LeuGlnLeuLeuLysGlnSerSerArgArgLeuGlnHisThrPheValPheLeu----- 30
Db 34 ATGCAGTTTTCGGGAAACCTCCACTCAGCTTCAACATCTTATTATTTATTTATTTAC 93
Qy 31 ArgAenPheSerLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50
Db 94 CAAATTTT-----ATTATTAGTTCTTAGTCAACGGGCGCTCCACGACATTC 144

RESULT 36
US-09-134-000C-3042/C
; Sequence 3042, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3042
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3042

Alignment Scores:
Pred. No.: 8.3 Length: 504
Score: 60.50 Matches: 13
Percent Similarity: 71.43% Conservative: 7
Best Local Similarity: 46.43% Mismatches: 7
Query Match: 19.71% Indels: 1
DB: 4 Gaps: 1

US-10-031-158-14 (1-58) x US-09-134-000C-3042 (1-504)

Qy 8 ProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSerArgArgLeuGlnHisThr---- 26
Db 370 CATTGTTTTCGGGCTGGATCTGATTATTCAAAGCGGCGCGTTGGCAACATTTCACAC 311
Qy 27 PheValPheLeuArgAsnPheSer 34
Db 310 TTTCATATATCTGTCAACGTTTTCG 287

```

RESULT 37
US-08-943-731-7/c
; Sequence 7, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILIA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; STREET: FLR.
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY Ph.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-943-731-7

Alignment Scores:
Pred. No.: 37.5 Length: 1463
Score: 60.50 Matches: 17
Percent Similarity: 42.86% Conservative: 4
Best Local Similarity: 34.69% Mismatches: 15
Query Match: 19.71% Indels: 13
DB: 3 Gaps: 1

US-10-031-158-14 (1-58) x US-08-943-731-7 (1-1463)
QY 4 PhePProPSePProLeuPhePheLeuGlnLeuLeuGlnSerSerArgArgLeu 23
DB 1366 TTCGCCGCCCTCCGCCGTTCTTCTTCAAAATCCTTAAAGCTCGCCTCCTCAT 1307
QY 24 GIWHLS-----ThrPheValPheLeu 30

DB 1306 CAGCGCGGTGACAGCGCGCGCTGCGCAGGAGCTGTACTCTTTCAGTCTC 1247
QY 3i ArgAnPheSerLeuMetLeuLeuArg 39
DB 1246 AGGAATTTAAACAAAGCTTGTGTCGC 1220

RESULT 38
US-08-943-731-1/c
; Sequence 1, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILIA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; STREET: FLR.
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY Ph.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18609 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-943-731-1

Alignment Scores:
Pred. No.: 1.38e+03 Length: 18609
Score: 60.50 Matches: 17
Percent Similarity: 42.86% Conservative: 4
Best Local Similarity: 34.69% Mismatches: 15
Query Match: 19.71% Indels: 13
DB: 3 Gaps: 1


```

US-10-031-158-14 (1-56) x US-08-943-731-1 (1-18609)
Qy      4 PheProPserProLeuPhePheLeuGlnLeuLeuYgInSerSerArgArgLeu 23
Db      1888 TTCCCGCCGCCCTCCGCCGCTCTCTTCTTCTCAATCTTAAAGCTGCCTGCTCAT 1829
Qy      24 GlnHis-----ThrPheValPheLeu 30
Db      1828 CAGCGCGGGTACACAGCGCCGAGCGCTGGCCAGGAGGCTGTAACTTTTCAGATTCTC 1769
Qy      31 ArgAsnPheSerLeuMetLeuLeuArg 39
Db      1768 AGCAATTAAACAAGCTTTAGTCCGC 1742

RESULT 39
US-09-543-681A-392
/ Sequence 392, Application US/09543681A
/ Patent No. 6605709
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/09/543,681A
/ CURRENT FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 392
/ LENGTH: 2715
/ TYPE: DNA
/ ORGANISM: Proteus mirabilis
US-09-543-681A-392

Alignment Scores:
Pred. No.:      107      Length:      2715
Score:          60.00      Matches:      18
Percent Similarity: 49.02% Conservative: 7
Best Local Similarity: 35.29% Mismatches: 16
Query Match:      19.54% Indels:      10
DB:               4      Gaps:      2

US-10-031-158-14 (1-56) x US-09-543-681A-392 (1-2715)
Qy      15 LeuLeuYgInSerSerArgArgLeuGlnIsthrPheValPheLeuArgAsnPheSer 34
Db      610 CTGGAGAGACAGACAGAGCGTGCATATAGAGCAT-----TTTTCGTAAAGCTAAT 660
Qy      35 LeuMetLeuLeuArgTyrIleGlyLylySerArgAlaThrArg-----49
Db      661 TTGATTGCAATTCGAGAGGCTTCCCTTAACCAATGGCAGATGCGTTGATTCGCAAGTA 720
Qy      50 -----PheTyrAspProArgArgGlyThrPro 58
Db      721 AAAGAATTCGTGATAGCCCAAGGTGTACGCT 753

RESULT 40
US-09-534-638-1
/ Sequence 1, Application US/09543638
/ Patent No. 6320038
/ GENERAL INFORMATION:
/ APPLICANT: Panula, Pertti A.J.
/ APPLICANT: Brandt, Annika
/ APPLICANT: Westerlund, Johanna
/ TITLE OF INVENTION: Promoter for Neuropeptide FF Promoter and use thereof
/ TITLE OF INVENTION: for therapy and diagnosis
/ FILE REFERENCE: 2530-104
/ CURRENT APPLICATION NUMBER: US/09/534,638
/ CURRENT FILING DATE: 2000-03-27
/ EARLIER APPLICATION NUMBER: 09/365755
/ EARLIER FILING DATE: 1999-08-03
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn Ver. 2.1

```

```

: SEQ ID NO 1
: LENGTH: 9840
: TYPE: DNA
: ORGANISM: Mouse
US-09-534-638-1

Alignment Scores:
Pred. No.: 662 Length: 9840
Score: 60.00 Matches: 18
Percent Similarity: 39.06% Conservative: 7
Best Local Similarity: 28.12% Mismatches: 27
Query Match: 19.54% Indels: 12
Gaps: 1

US-10-031-158-14 (1-58) x US-09-534-638-1 (1-9840)

QY      2 GlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSerArg 21
Db      1041 CAGGGTGTCCACCTCTTCATGTGCTTATTCCTATCTGCTACCAAAACAAACACAAA 1100
        ::::
QY      22 ArgLeuGlnHisThrPheValPheLeuArgAsnPheSer-----34
Db      1101 CAAAAAAGCAAAACTCGAGAGTTGATCATTTTCATTCCAAATCTATTGGTACT 1166
        ::::
QY      35 -----LeuMetLeuLeuArgTyrIleGlyLysLysAlaGArgAlaThrArg 49
Db      1161 CATTTGCTCTGTGACTTTCTTCTAGTCTAGGAGTGAGGTCTTCTCTACTCTCCCACTAGG 1220
        ::::
QY      50 PheTrpAspPro 53
        |||::|||
Db      1221 CTTGGGAACT 1232

```

Search completed: December 8, 2004, 09:20:25
Job time : 71 secs

Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2004, 08:54:21 ; Search time 271 Seconds

(without alignments)
1180.016 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307
Sequence: 1 MOMPSPPLFFELQLLKSS.....RYGKRRATRPWDRGRGP 58

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=x1h
-O=/cgn2_1/USPTO/US10031158/runat_06122004_082701_16006/app_query.fasta_1.199
-DB=Published Applications NA -QMT=fastcap -SUFFIX=trpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blousum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=0 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINTEN=0
-MAXLEN=200000000 -USPR=US10031158_@CGN_1.1.354_@runat_06122004_082701_16006
-ICPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	470	9	US-09-864-761-15428
2	307	100.0	477	9	US-09-864-761-15438
3	307	100.0	825	15	US-10-340-536-29
4	307	100.0	1027	15	US-10-205-823-40
5	307	100.0	1395	18	US-10-643-755A-23
6	307	100.0	1418	10	US-09-957-708-16
7	302	98.4	316	9	US-09-864-761-31943
8	294	95.8	1421	9	US-09-954-456-317
9	294	95.8	1799	15	US-10-101-510-492
10	284	92.5	1586	10	US-09-960-706-676
11	284	92.5	1586	10	US-09-873-319-424
12	245	79.8	1155	9	US-09-925-300-58
13	75	24.4	1813	17	US-10-437-963-35632
14	73.5	23.9	618	13	US-10-027-632-136731
15	73.5	23.9	618	13	US-10-027-632-136731
16	72.5	23.6	618	13	US-10-027-632-105504
17	72.5	23.6	618	13	US-10-027-632-105505
18	72.5	23.6	618	13	US-10-027-632-105505
19	72.5	23.6	618	15	US-10-027-632-105505
20	72.5	23.6	618	15	US-09-764-891-10230
21	72.5	23.6	27118	15	US-10-205-428-1018
22	71.5	23.3	328	16	US-10-424-599-63943
23	70.5	23.0	618	13	US-10-027-632-136730
24	70.5	22.0	618	15	US-10-027-632-136730
25	69.5	22.6	353	16	US-10-424-599-127254
26	69.5	22.6	2145	16	US-10-282-122A-22460
27	69.5	22.6	2241	9	US-09-881-752A-321
28	68.5	22.3	585	13	US-10-027-632-203503
29	68.5	22.3	585	18	US-10-027-632-203503
30	67.5	22.0	535	18	US-10-363-345A-15079
31	67.5	22.0	535	18	US-10-363-345A-15080
32	67.5	22.0	942	9	US-09-962-833-256
33	67.5	22.0	1037	15	US-10-172-118-1338
34	67.5	22.0	1037	16	US-10-342-887-1338
35	67.5	22.0	2532	10	US-09-918-648-48
36	67.5	22.0	11838	10	US-09-764-891-10230
37	67.5	22.0	13814	15	US-10-311-455-1165
38	67	21.8	314	18	US-10-425-115-4936
39	67	21.8	1265	13	US-10-027-632-252580
40	67	21.8	1265	15	US-10-027-632-252580
41	67	21.8	1726	18	US-10-425-115-128054
42	67	21.8	29829	13	US-10-087-192-694
43	67	21.8	3673778	15	US-10-312-841-2
44	66.5	21.7	328	18	US-10-424-599-134189
45	66.5	21.7	1316	18	US-10-425-115-135105

ALIGNMENTS

RESULT 1
US-09-864-761-15428
; Sequence 15428, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6

```
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 15428
/ LENGTH: 470
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AF159056.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
US-09-864-761-15428

Alignment Scores:
Pred. No.: 5,1e-36 Length: 470
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-15428 (1-470)

QY 1 MetGlnMetPheProPsePProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
DB 203 ATGCAAGTGTTCCTCCCAAGCCACCTATTTCTTCCTCAATTCGGAACAAGCTCC 262

QY 21 ArgArgLeuGlnIstIstIstIstIstIstIstIstIstIstIstIstIstIstIst 40
DB 263 AGAAGGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGAGTTATTAAAGATAC 322

QY 41 IIGGLVLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSL 58
DB 323 ATTGGCAAGAAAGAAAGAAAGCAACGATTCGGATCCAGAGGGAGCAACACCA 376

RESULT 2
US-09-864-761-15438
/ Sequence 15438, Application US/09864761
/ Patient No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aeomica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 15438
/ LENGTH: 477
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AF159056.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
US-09-864-761-15438

Alignment Scores:
Pred. No.: 5,21e-36 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-15438 (1-477)

QY 1 MetGlnMetPheProPsePProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
DB 212 ATGCAAGTGTTCCTCCCAAGCCACCTATTTCTTCCTCAATTCGGAACAAGCTCC 271

QY 21 ArgArgLeuGlnIstIstIstIstIstIstIstIstIstIstIstIstIstIstIst 40
DB 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGAGTTATTAAAGATAC 331

QY 41 IIGGLVLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSL 58
```

```

P R I O R F I L I N G D A T E : 2001-08-22
P R I O R A P P L I C A T I O N N U M B E R : 60/325,020
P R I O R F I L I N G D A T E : 2001-09-25
P R I O R A P P L I C A T I O N N U M B E R : 60/341,746
P R I O R F I L I N G D A T E : 2001-12-12
P R I O R A P P L I C A T I O N N U M B E R : 60/362,158
P R I O R F I L I N G D A T E : 2002-03-05
N U M B E R O F S E Q I D N O S : 455
S O F T W A R E : PaSeQ for Windows Version 4.0
S E Q I D N O 404
L E N G T H : 1027
T Y P E : DNA
O R G A N I S M : Homo sapiens
U S -10-205-823-404

Alignment Scores:
Pred. No.:      1,5e-35      Length:      1027
Score:          307.00      Matches:     58
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:      100.00%     Indels:      0
DB:               15         Gaps:        0

US-10-031-158-14 (1-58) x US-10-205-823-404 (1-1027)

Cy       1 MetGlnMetPheProProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
Db       74 ATGCACATGTTCCTCCCAAGCCCAAGCCACTATTTCCTTCCTCAATGGCGAAACAAGCTCC 133

Cy       21 ArgArgLeuGlnHisThrPheValPheLeuArgAenPheSerLeuMetLeuLeuArgTyr 40
Db       134 AGAGAGCTGAACARDCCTTTTCCTTCCTTGAGAAATTTTTCCCTGATGTTATTAAAGATAC 199

Cy       41 IleGlyLeuAlaArgArgIaThrArgPheTPaApproArgArgGlyThrPro 58
Db       194 ATTGGCAGAAGAAAAGAACAGCACAGATTCCTGGATCCAGAGGGGAACACCA 247

RESULT 5
US-10-643-795A-23
; Sequence 23, Application US/10643795A
; Publication No. US20040241703A1
GENERAL INFORMATION:
APPLICANT: FREDERIC J. DESAUVAGE
APPLICANT: GRETCHEN FRANTZ
APPLICANT: KENNETH J. HILLAN
APPLICANT: PAUL POLAKIS
APPLICANT: ANDREW POLSON
APPLICANT: VICTORIA SMITH
APPLICANT: SUSAN D. SPENCER
APPLICANT: THOMAS D. WU
APPLICANT: ZEMIN ZHANG
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
TITLE OR INVENTION: TREATMENT OF TUMOR
FILE REFERENCE: P5026R1-US
CURRENT APPLICATION NUMBER: US/10/643,795A
CURRENT FILING DATE: 2003-08-19
PRIOR APPLICATION NUMBER: US 60/404,809
PRIOR FILING DATE: 2002-08-19
PRIOR APPLICATION NUMBER: US 60/405,645
PRIOR FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: US 60/413,192
PRIOR FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: US 60/419,008
PRIOR FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: US 60/426,847
PRIOR FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US 60/484,959
PRIOR FILING DATE: 2003-07-02
NUMBER OF SEQ ID NOS: 158
SEQ ID NO 23
LENGTH: 1395
TYPE: DNA
ORGANISM: Homo sapien

```

US-10-643-795A-23

Alignment Scores:

Pred. No.:	2,296-35	Length:	1395
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	18	Gaps:	0

US-10-031-158-14 (1-58) x US-10-643-795A-23 (1-1395)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
DB 449 ATGCAGATGTTTCCCCCAAGCCACATATTTCTTCTTCAATGCTGAAACAAAGCTCC 508

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuLeuLeuArgTyr 40
DB 509 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 568

QY 41 IleGlyValPheArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
DB 569 ATTGGCAAGAAAGAGACACACATTTCTGGATCCAGAGGGGAACACCA 622

RESULT 6

US-09-957-708-16
Sequence 16, Application US/09957708
Publication No. US20030031678A1
GENERAL INFORMATION:
APPLICANT: Sun, Yongming
APPLICANT: Recipon, Heive
APPLICANT: Cafetkey, Robert
APPLICANT: Ali, Shujah
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
FILE REFERENCE: DEX-0239
CURRENT APPLICATION NUMBER: US/09/957,708
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/233,746
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 1418
TYPE: DNA
ORGANISM: Homo sapiens
US-09-957-708-16

Alignment Scores:

Pred. No.:	2,356-35	Length:	1418
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-031-158-14 (1-58) x US-09-957-708-16 (1-1418)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20
DB 449 ATGCAGATGTTTCCCCCAAGCCACATATTTCTTCTTCAATGCTGAAACAAAGCTCC 508

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuLeuLeuArgTyr 40
DB 509 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 568

QY 41 IleGlyValPheArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
DB 569 ATTGGCAAGAAAGAGACACACATTTCTGGATCCAGAGGGGAACACCA 622

RESULT 7

US-09-864-761-31943
Sequence 31943, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 31943
LENGTH: 316
TYPE: DNA
ORGANISM: Homo sapiens

FEATURES:

OTHER INFORMATION: MAP TO AF159056.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
OTHER INFORMATION: SWISSPROT HIT: P03986, EVALUE 7.00e-58
OTHER INFORMATION: NT HIT: M44996.1, EVALUE 0.00e+00
OTHER INFORMATION: EST_HUMAN HIT: BF679123.1, EVALUE 0.00e+00
US-09-864-761-31943

Alignment Scores:

Pred. No.:	1,646-35	Length:	316
Score:	302.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.37%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x US-09-864-761-31943 (1-316)

```
QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSerArg 21
| | | | |
Db 2 CAGAGTGTTCCTCCCAAGCCACATTTTCTTCTTCATTCGTGAACAAAGCTCCGA 61
QY 22 ArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyrIle 41
| | | | |
Db 62 AGGCTGGAACATACATCTTCTTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATACATT 121
QY 42 GLyLyLyLyArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
| | | | |
Db 122 GGCAGAGAAAGAGAGACACACGATTCGAGATCCAGAGGAGGAGAACACCA 172
RESULT 8
US-09-954-456-317
; Sequence 317, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 317
; TYPE: DNA
; LENGTH: 1421
; ORGANISM: Homo sapiens
US-09-954-456-317
Alignment Scores:
Pred. No.: 2,056-33 Length: 1421
Score: 294.00 Matches: 55
Percent Similarity: 98.28% Conservative: 2
Best Local Similarity: 94.83% Mismatches: 1
Query Match: 95.77% Indels: 0
Gaps: 0
DB: 9
US-10-031-158-14 (1-58) x US-09-954-456-317 (1-1421)
QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
| | | | |
Db 423 ATGCAATATTTTCCCAAGCCACATTTTCTTCTTCATTCGTGAACAAAGCTCC 482
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
| | | | |
Db 483 AGAAGGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCAGATATTATTAAATAC 542
QY 41 ILeGlyLyLyArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
| | | | |
Db 543 ATTGGCAAGAAAGAGAGACACGATTCGAGATCCAGAGGAGGAGAACACCA 596
```

```
RESULT 9
US-10-101-510-492
; Sequence 492, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117,0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 492
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-492
Alignment Scores:
Pred. No.: 2,846-33 Length: 1799
Score: 294.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 95.77% Mismatches: 1
Query Match: 95.77% Indels: 0
Gaps: 0
DB: 15
US-10-031-158-14 (1-58) x US-10-101-510-492 (1-1799)
QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
| | | | |
Db 845 ATGCAATATTTTCCCAAGCCACATTTTCTTCTTCATTCGTGAACAAAGCTCC 904
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
| | | | |
Db 905 AGAAGGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATAC 964
QY 41 ILeGlyLyLyArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
| | | | |
Db 965 ATTGGCAAGAAAGAGAGACACGATTCGAGATCCAGAGGAGGAGAACACCA 1019
RESULT 10
US-09-960-706-676
; Sequence 676, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William B.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M30894
US-09-960-706-676
Alignment Scores:
Pred. No.: 7,436-32 Length: 1586
Score: 284.00 Matches: 53
Percent Similarity: 96.55% Conservative: 3
Best Local Similarity: 91.38% Mismatches: 2
```

```
Query Match: 92.51% Indels: 0
DB: 10 Gaps: 0
US-10-031-158-14 (1-58) x US-09-960-706-676 (1-1586)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
DB 553 ATGCAGATGTTTCCCAAGCCACATATTTTCTTCCTTCGATTCGAAACAACTCC 612
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
DB 613 AGAAGGCTGGAACATATCTTTGCTTCTTGAGAAATTTTCCAGATATTTAAGATAC 672
QY 41 lIeGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
DB 673 ATTGCAGAAAGAAAGAGCAACACGATTCGGGATCCCGAGGGGAAACACCA 726

RESULT 11
US-09-873-319-424
; Sequence 424, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William F.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Mago, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; EARLIER FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 424
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M30894
US-09-873-319-424

Alignment Scores:
Pred. No.: 7,43e-32 Length: 1586
Score: 284.00 Matches: 53
Percent Similarity: 96.55% Conservative: 3
Best Local Similarity: 91.38% Mismatches: 2
Query Match: 92.51% Indels: 0
Gaps: 0

US-10-031-158-14 (1-58) x US-09-873-319-424 (1-1586)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
DB 553 ATGCAGATGTTTCCCAAGCCACATATTTTCTTCCTTCGATTCGAAACAACTCC 612
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
DB 613 AGAAGGCTGGAACATATCTTTGCTTCTTGAGAAATTTTCCAGATATTTAAGATAC 672
QY 41 lIeGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
DB 673 ATTGCAGAAAGAAAGAGCAACACGATTCGGGATCCCGAGGGGAAACACCA 726

RESULT 12
US-09-925-300-58
; Sequence 58, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (432)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (443)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-58

Alignment Scores:
Pred. No.: 3.19e-26 Length: 1155
Score: 245.00 Matches: 51
Percent Similarity: 91.07% Conservative: 0
Best Local Similarity: 91.07% Mismatches: 5
Query Match: 79.80% Indels: 1
Gaps: 0

US-10-031-158-14 (1-58) x US-09-925-300-58 (1-1155)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
DB 175 ATGCAGATGTTTCCCAAGCCACATATTTCTTCCTTCGATTCGAAACAACTCC 233
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
DB 234 AGAAGGCTGGAACATATCTTTGCTTCTTGAGAAATTTTCCAGATATTTAAGATAC 293
QY 41 lIeGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGly 56
DB 294 ATTGSBAAGAAAGAGCAACACGATTCGGGATCCCGAGGGG 341

RESULT 13
US-10-437-963-35632/C
; Sequence 35632, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 35632
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_39533C.1
```


TITLE OF INVENTION: Polymorphisms in the Human Genome

TITLE OF INVENTION: Polymorphisms in the Human Genome


```
/ Sequence 321, Application US/09881752A
/ Patent No. US20020115078A1
/ GENERAL INFORMATION:
/ APPLICANT: Kleinhous, Harold
/ APPLICANT: Al-Garawi, Amal
/ APPLICANT: Miller, Charles
/ APPLICANT: Tomb, Jean-Francois
/ APPLICANT: Oomen, Raymond P.
/ TITLE OF INVENTION: Identification of Polynucleotides
/ TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
/ TITLE OF INVENTION: Genome
/ FILE REFERENCE: 06132/041002
/ CURRENT APPLICATION NUMBER: US/09/881,752A
/ CURRENT FILING DATE: 2001-06-15
/ PRIOR APPLICATION NUMBER: US 08/833,457
/ PRIOR FILING DATE: 1997-04-01
/ NUMBER OF SEQ ID NOS: 370
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 321
/ LENGTH: 2241
/ TYPE: DNA
/ ORGANISM: Helicobacter pylori
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (52)...(2193)
/ US-09-881-752A-321

Alignment Scores:
Pred. No.: 12.7 Length: 2241
Score: 69.50 Matches: 25
Percent Similarity: 51.61% Conservative: 17
Best Local Similarity: 40.32% Mismatches: 7
Query Match: 22.64% Indels: 14
DB: 9 Gaps: 3

US-10-031-158-14 (1-58) x US-09-881-752A-321 (1-2241)

QY 1 MetGlnMetPheProSerPro-----Leu 9
Db 2130 ATACAAATGTTCCCTCCCTGACTTCCTCATCTTAAACGCAATTCTAGCCTT 2071
QY 10 PhePhePheGlnLeuLeuLeuGlnSerSerArgArgLeuGlnHisThrPheValPhe 29
Db 2070 GATATTTTCTTGCAA--TTCAAAAGCAGAGGCGCATGTGACCTTCAATCTGCTTT 2014
QY 30 LeuArgAsnHisSerLeuMetLeuLeuArgTyrIleGlyValysArgAlaIleThrArg 49
Db 2013 TCCAGGCT-TTT--CTAGTTTAAAAATCGTTTCAAAATCTAGCGCTGCTTCTT 1958
QY 50 PheTyr 51
Db 1957 TTTTGG 1952

RESULT 28
US-10-027-632-203503
/ Sequence 203503, Application US/10027632
/ Publication No. US2002019837A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/199,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
```

```
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 203503
/ LENGTH: 585
/ TYPE: DNA
/ ORGANISM: Human
/ US-10-027-632-203503

Alignment Scores:
Pred. No.: 2.79 Length: 585
Score: 68.50 Matches: 18
Percent Similarity: 49.09% Conservative: 9
Best Local Similarity: 32.73% Mismatches: 19
Query Match: 22.31% Indels: 9
DB: 13 Gaps: 2

US-10-031-158-14 (1-58) x US-10-027-632-203503 (1-585)

QY 5 ProProSerProLeuPhePhePheGlnLeuLeuGlnSerSerArgArgLeuGlu 24
Db 361 CCTCCCTTCCCTCCCTCCACTTTTC-----CAATCTCTTCTCTTATCC 405
QY 25 HisThrPheValPheLeuArgAsnHisSerLeuMetLeu-----ArgTyr 40
Db 406 CTAACTTATTTATTTATTTTGGATTCTTATCTATCTGTACACACTTGACATTAAT 465
QY 41 IleGlyValysArgArgAlaThrArgPheTyrPaspProArgArg 55
Db 466 GTAGGAAAAAGAGCTGGCATTTTCTTATTTGTAAACCAAGAAA 510

RESULT 29
US-10-027-632-203503
/ Sequence 203503, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 203503
/ LENGTH: 585
/ TYPE: DNA
/ ORGANISM: Human
/ US-10-027-632-203503

Alignment Scores:
Pred. No.: 2.79 Length: 585
Score: 68.50 Matches: 18
Percent Similarity: 49.09% Conservative: 9
Best Local Similarity: 32.73% Mismatches: 19
```

```

CURRENT APPLICATION NUMBER: US/10/363,345A
CURRENT FILING DATE: 2003-03-03
NUMBER OF SEQ ID NOS: 40712
SEQ ID NO 15080
LENGTH: 535
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-363-345A-15080

Alignment Scores:
Pred. No.: 3.48 Length: 535
Score: 67.50 Matches: 18
Percent Similarity: 55.56% Conservative: 7
Best Local Similarity: 40.00% Mismatches: 15
Query Match: 21.99% Indels: 5
DB: 18 Gaps: 2

US-10-031-158-14 (1-58) x US-10-363-345A-15080 (1-535)
QY 12 PheLeuGlnLeuLeuLysGlnSerSerArgArgLeuGlnHisThrPheValPheLeuArg 31
Db 454 TTTTTCGATCGCGAAGCGG-----CGGCGTTTACTTTTCGTTTTTTTTTTGTCGG 401
QY 32 AsnPheSerLeuWetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPheTrp 51
Db 400 GTTTCGCGTAAATTTTATTATTGTTGGTGGATCGGTTTCGGGTTCTGAGCTTT-- 344
QY 52 AspProArgArgGly 56
Db 343 -----CGTCGGCGA 335

RESULT 32
US-09-962-832-256
Sequence 256, Application US/09962832
Patent No. US20020110821A1
GENERAL INFORMATION:
APPLICANT: Ebner, Reinhard
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-74
CURRENT APPLICATION NUMBER: US/09/962,832
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/60/235,077
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,280
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 259
SOFTWARE: PatentIn version 3.0
SEQ ID NO 256
LENGTH: 942
TYPE: DNA
ORGANISM: Homo sapiens
US-09-962-832-256

Alignment Scores:
Pred. No.: 7.6 Length: 942
Score: 67.50 Matches: 17
Percent Similarity: 61.11% Conservative: 5
Best Local Similarity: 47.22% Mismatches: 11
Query Match: 21.99% Indels: 3
DB: 9 Gaps: 2

US-10-031-158-14 (1-58) x US-09-962-832-256 (1-942)
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuWetLeu--Arg 39
Db 620 AGACATCTGCTGCATACCGCTACTATTTGATGATGATTTTCCATGCCCTTGATACGGAGA 679
QY 40 TyrIleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArg 55
Db 400 GTTTCGCGTAAATTTTATTATTGTTGGTGGATCGGTTTCGGGTTCTGAGCTTT-- 344

```


Db 680 TACGTAGAGAAAGAAAGAGG-----CAGTACTGGCCATGAGAGAAA 721

RESULT 33

US-10-172-118-1338

Sequence 1338, Application US/10172118

Publication No. US20030224374A1

GENERAL INFORMATION:

APPLICANT: Dai, Hongyue

APPLICANT: He, Yudong

APPLICANT: Linsley, Peter

APPLICANT: Mao, Mao

APPLICANT: Roberts, Chris

APPLICANT: Van 't Veer, Laura

APPLICANT: Van de Vijver, Marc

APPLICANT: Bernardes, Rene

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-175-999

CURRENT APPLICATION NUMBER: US/10/172,118

CURRENT FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 60/380,770

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 2659

SEQ ID NO 1338

LENGTH: 1037

TYPE: DNA

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: NM_006912

DATABASE ENTRY DATE: 2001-06-18

US-10-172-118-1338

Alignment Scores:

Pred. No.:	Score:	Length:
8.68	67.50	1037
61.11%	Conservative: 5	Matches: 17
47.22%	Mismatches: 11	
21.99%	Indels: 3	
	Gaps: 2	

US-10-031-158-14 (1-58) x US-10-172-118-1338 (1-1037)

QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeu---Arg 39

Db 620 AGACATCTGCTGCATACCGCTACTATATGATGATGATTTTCCATGCCCTTGTAACGGAGAGA 679

QY 40 TyrTleGlyLysLysArgAlaThrArgPheTrpAspProArgArg 55

Db 680 TACGTAGAGAAAGAAAGAGG-----CAGTACTGGCCATGAGAGAAA 721

RESULT 34

US-10-342-887-1338

Sequence 1338, Application US/10342887

Publication No. US20040058340A1

GENERAL INFORMATION:

APPLICANT: Dai, Hongyue

APPLICANT: He, Yudong

APPLICANT: Linsley, Peter S.

APPLICANT: Mao, Mao

APPLICANT: Roberts, Christopher J.

APPLICANT: Van 't Veer, Laura Johanna

APPLICANT: Van de Vijver, Marc J.

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-188-999

CURRENT APPLICATION NUMBER: US/10/342,887

CURRENT FILING DATE: 2003-01-15

PRIOR APPLICATION NUMBER: 60/298,918

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/380,720

PRIOR FILING DATE: 2002-05-14

PRIOR APPLICATION NUMBER: 10/172,118

PRIOR FILING DATE: 2002-06-14

NUMBER OF SEQ ID NOS: 2659

SEQ ID NO 1338

LENGTH: 1037

TYPE: DNA

ORGANISM: Homo sapiens

US-10-342-887-1338

Alignment Scores:

Pred. No.:	Score:	Length:
8.68	67.50	1037
61.11%	Conservative: 5	Matches: 17
47.22%	Mismatches: 11	
21.99%	Indels: 3	
	Gaps: 2	

US-10-031-158-14 (1-58) x US-10-342-887-1338 (1-1037)

QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeu---Arg 39

Db 620 AGACATCTGCTGCATACCGCTACTATATGATGATGATTTTCCATGCCCTTGTAACGGAGAGA 679

QY 40 TyrTleGlyLysLysArgAlaThrArgPheTrpAspProArgArg 55

Db 680 TACGTAGAGAAAGAAAGAGG-----CAGTACTGGCCATGAGAGAAA 721

RESULT 35

US-09-918-624B-4

Sequence 4, Application US/0918624B

Publication No. US20030113720A1

GENERAL INFORMATION:

APPLICANT: Schebye, Xiao Min

APPLICANT: Sornasse, Thierly

TITLE OF INVENTION: CDNAS EXPRESSED IN ADIPOCYTE DIFFERENTIATION

FILE REFERENCE: PA-0033 US

CURRENT APPLICATION NUMBER: US/09/918,624B

CURRENT FILING DATE: 2002-12-03

PRIOR APPLICATION NUMBER: 60/222,470

PRIOR FILING DATE: 2000-07-28

NUMBER OF SEQ ID NOS: 71

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 2532

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20030113720A1 481536.3

US-09-918-624B-4

Alignment Scores:

Pred. No.:	Score:	Length:
29.8	67.50	2532
61.11%	Conservative: 5	Matches: 17
47.22%	Mismatches: 11	
21.99%	Indels: 3	
	Gaps: 2	

US-10-031-158-14 (1-58) x US-09-918-624B-4 (1-2532)

QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeu---Arg 39

Db 732 AGACATCTGCTGCATACCGCTACTATATGATGATGATTTTCCATGCCCTTGTAACGGAGAGA 791

QY 40 TyrTleGlyLysLysArgAlaThrArgPheTrpAspProArgArg 55

Db 792 TACGTAGAGAAAGAAAGAGG-----CAGTACTGGCCATGAGAGAAA 833

RESULT 36

US-09-764-891-7020

Sequence 7020, Application US/09764891

Publication No. US20030077808A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7020
LENGTH: 11838
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-7020

Alignment Scores:
Pred. No.: 251          Length: 11838
Score: 67.50           Matches: 17
Percent Similarity: 61.1% Conservative: 5
Best Local Similarity: 47.22% Mismatches: 11
Query Match: 21.99% Indels: 3
DB: 10 Gaps: 2

US-10-031-158-14 (1-58) x US-09-764-891-7020 (1-11838)
QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPhseSerLeuMetLeu---Arg 39
Db 10840 AGACATCTGCTGCATACCGCTACTATATTCATGATGTTTCCATGCGCTTGTCGGGAGA 10899
QY 40 TyrTlGlyLylslyArgAlaThrArgPheTyrAspProArgArg 55
Db 10900 TACGTAGAAAGAAAGAGAG-----CAGTACTGCACATGAGAGAA 10941

RESULT 37
US-10-311-455-1165
Sequence 1165, Application US/10311455
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of Specific Genes
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1165
LENGTH: 13814
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
NAME/KEY: unsure
LOCATION: 4518
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1165

Alignment Scores:
Pred. No.: 311          Length: 13814
Score: 67.50           Matches: 18
Percent Similarity: 55.56% Conservative: 7
Best Local Similarity: 40.00% Mismatches: 15
Query Match: 21.99% Indels: 5
DB: 15 Gaps: 2

US-10-031-158-14 (1-58) x US-10-311-455-1165 (1-13814)
QY 12 PheLeuGlnLeuLeuysGlnSerSerArgArgLeuGluHisThrPheValPheLeuArg 31
```

```
Db 5345 TTTTTCATGCTGAAGCG-----CGCGTTTACTTTGCTTTTTCGTCGG 5398
QY 32 AsnPhseSerLeuMetLeuLysGlyTyrTlGlyLylslyArgAlaThrArgPheTyr 51
Db 5399 GTTTCGCGTATTTTATTTTATTTGTTGGATCGCGTTTCGGGTTTCGTAGTTT--- 5455
QY 52 AspProArgArgGly 56
Db 5456 -----CGTCGGGGA 5464

RESULT 38
US-10-425-115-4936/c
Sequence 4936, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants
FILE REFERENCE: 38-21(5322)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 4936
LENGTH: 314
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_104489C.1
US-10-425-115-4936

Alignment Scores:
Pred. No.: 1.98          Length: 314
Score: 67.00           Matches: 20
Percent Similarity: 45.76% Conservative: 7
Best Local Similarity: 33.90% Mismatches: 20
Query Match: 21.82% Indels: 12
DB: 18 Gaps: 3

US-10-031-158-14 (1-58) x US-10-425-115-4936 (1-314)
QY 5 ProProSerProLeuPhePheLeuGln-----LeuLysGlnSerSerArg 21
Db 311 CCCCCACCCCGACGATTTTTCATGCCCAATGCCCTTTCTTGCGTTTGCCAA 252
QY 22 ArgLeuGluHisThrPheValPheLeuArgAsn-----PheSerLeuMetLeuArg 39
Db 251 AGGCTTCCCGCTCAATCTTTTATAGGGGCCCACTTTAGATTGCTCCTT--- 198
QY 40 TyrTlGlyLylslyArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
Db 197 -----AGACCCCGCAGAACTTTTCCCAAGAAAAATTGCCCC 156

RESULT 39
US-10-027-632-252580/c
Sequence 252580, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:33 ; Search time 181 Seconds

(without alignments)
140.795 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58
Sequence: 1 MOWPSPPLFFPLQLKQSS.....RYIGKKRRARFMDPRRGRP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 243163 seqs, 439378781 residues

Word size: 0

Total number of hits satisfying chosen parameters: 408321

Minimum DB seq length: 8
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database:

1: Genesegp21:*
2: genesegp1980s:*
3: genesegp2000s:*
4: genesegp2001s:*
5: genesegp2002s:*
6: genesegp2003as:*
7: genesegp2003bs:*
8: genesegp2004s:*
9: genesegp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	15.5	9	ADM44402	Adm44402 Human TAR
2	9	15.5	9	ADM44401	Adm44401 Human TAR
3	9	15.5	9	ADM44406	Adm44406 Human TAR
4	8	13.8	8	ADM44405	Adm44405 Human TAR
5	8	13.8	9	ADM44404	Adm44404 Human TAR
6	6	10.3	9	ADH89809	Adh89809 Cell gene
7	6	10.3	9	ADM44403	Adm44403 Human TAR
8	5	8.6	8	AAB06282	Aab06282 p10SKI-1
9	5	8.6	8	AEA41508	Aea41508 GAG-blndi
10	5	8.6	9	AAR59128	Aar59128 Peptide f
11	5	8.6	9	AAV38039	Aav38039 Hepatitis
12	5	8.6	9	AAV38106	Aav38106 Hepatitis
13	5	8.6	9	AAW44612	AAw44612 Anti-fung
14	5	8.6	9	AAW43777	AAw43777 Bacterici
15	5	8.6	9	AAV45608	AAv45608 Immunogen
16	5	8.6	9	AAV45674	AAv45674 Immunogen
17	5	8.6	9	AAV46768	AAv46768 Immunogen
18	5	8.6	9	AAV00589	AAv00589 Anti-funga
19	5	8.6	9	AAH65513	AAh65513 Anti-funga
20	5	8.6	9	AAU00796	AAu00796 HIV-1 Tat
21	5	8.6	9	ABB08361	ABb08361 Synthetic
22	5	8.6	9	AAO15791	AAo15791 Human imm
23	5	8.6	9	ABP54737	ABp54737 HIV-1 tat
24	5	8.6	9	AAE35004	AAe35004 Mouse imm

25	5	8.6	9	ABR82213	ABr82213 Human ant
26	5	8.6	9	ADK65211	ADk65211 Human PiG
27	5	8.6	9	ADK97590	ADk97590 Immunogen
28	5	8.6	9	ADH89810	ADh89810 Cell gene
29	5	8.6	9	ADH44385	ADh44385 INK4a imm
30	5	8.6	9	ADH44376	ADh44376 INK4a imm
31	5	8.6	9	ADK84035	ADk84035 Human 191
32	5	8.6	9	ADK84519	ADk84519 Human 191
33	5	8.6	9	ADK85026	ADk85026 Human 191
34	5	8.6	9	ADK85074	ADk85074 Human 191
35	5	8.6	9	ADK87550	ADk87550 Human 191
36	5	8.6	9	ADK88447	ADk88447 Human 191
37	5	8.6	9	ADK83950	ADk83950 Human 191
38	5	8.6	9	ADK88445	ADk88445 Human 191
39	5	8.6	9	ADK84560	ADk84560 Human 191
40	5	8.6	9	ADK85620	ADk85620 Human 191
41	5	8.6	9	ADK84518	ADk84518 Human 191
42	5	8.6	9	ADK86086	ADk86086 Human 191
43	5	8.6	9	ADK86641	ADk86641 Human 191
44	5	8.6	9	ADK87058	ADk87058 Human 191
45	5	8.6	9	ADK87549	ADk87549 Human 191

ALIGNMENTS

RESULT 1
ADM44402 standard; peptide; 9 AA.
XX
AC ADM44402;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human TARP polypeptide epitope TARP-27-35.
XX
KW T-cell receptor gamma alternate reading frame protein; TARP;
KW immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;
KW receptor.
XX
OS Homo sapiens.
XX
PN W0200500889-A1.
XX
PD 06-JAN-2005.
XX
PF 02-JUN-2004; 2004MO-US017574.
XX
PR 05-JUN-2003; 2003US-0476467P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Berzofsky JA, Oh S, Pastan I;
XX WPI; 2005-091494/10.
XX
XX New T-cell receptor gamma alternate reading frame protein (TARP)
XX polypeptide, useful for treating a subject having breast or prostate
XX cancer and for generating an immune response to TARP-expressing breast
XX and prostate cancer cells.
XX
XX Claim 2; SEQ ID NO 4; 83pp; English.
XX
XX The invention relates to an immunogenic T-cell receptor gamma alternate
XX reading frame protein (TARP) polypeptide and the polynucleotide encoding
XX it. The invention also relates to a vector comprising the TARP
XX polynucleotide, a host cell transformed with the vector, a pharmaceutical
XX composition comprising a therapeutic amount of the polypeptide or the
XX polynucleotide in a pharmaceutical carrier, eliciting an immune response
XX in a subject, inhibiting the growth of a breast cancer or a prostate
XX cancer cell, a reagent comprising a tetramer of the polypeptide bound to
XX HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
XX detecting T cells expressing CD8 that specifically recognize the TARP

CC polypeptide in a subject. Eliciting an immune response in a subject
CC comprises administering the polypeptide or polynucleotide, thus producing
CC an immune response in the subject. The immune response comprises a T cell
CC response or inducing cytotoxic T cells that induce lysis of cells
CC expressing the TARP polypeptide. The subject has breast cancer or
CC prostate cancer. The immune response decreases the growth of the prostate
CC cancer or breast cancer. The method further comprises administering an
CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
CC or CTL precursor cells with the polypeptide and an antigen presenting
CC cell to produce activated CTLs matured from the CTL precursors that
CC recognize the breast cancer or the prostate cancer cells, and contacting
CC the breast cancer or the prostate cancer cell with the activated CTLs or
CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
CC breast cancer or the prostate cancer cell. The TARP polypeptide and
CC polynucleotide are useful for treating a subject having breast or
CC prostate cancer. The polypeptide is useful for generating an immune
CC response to breast cancer and prostate cancer cells that express TARP
CC polypeptides. This sequence represents a human TARP polypeptide epitope
CC used in the scope of the invention.
CC
CC
SQ Sequence 9 AA;

Query Match 15.5%; Score 9; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 FVFLRNFSL 35
|||||

Db 1 FVFLRNFSL 9

RESULT 2
ADM44401
ID ADM44401 standard; peptide; 9 AA.

AC ADM44401;

DT 24-MAR-2005 (first entry)

DE Human TARP polypeptide epitope TARP-29-37.

XX T-cell receptor gamma alternate reading frame protein; TARP;
KW immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;
KW receptor.
XX

OS Homo sapiens.

FN WO200500889-A1.

PD 06-JAN-2005.

PF 02-JUN-2004; 2004WO-US017574.

PR 05-JUN-2003; 2003US-0476467P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Berzofsky JA, Oh S, Pastan I;

DR WPI; 2005-091494/10.

XX New T-cell receptor gamma alternate reading frame protein (TARP)
PT polypeptide, useful for treating a subject having breast or prostate
PT cancer and for generating an immune response to TARP-expressing breast
PT and prostate cancer cells.
XX

PS Claim 2; SEQ ID NO 3; 83bp; English.

XX The invention relates to an immunogenic T-cell receptor gamma alternate
CC reading frame protein (TARP) polypeptide and the polynucleotide encoding
CC it. The invention also relates to a vector comprising the TARP
CC polynucleotide, a host cell transformed with the vector, a pharmaceutical

CC composition comprising a therapeutic amount of the polypeptide or the
CC polynucleotide in a pharmaceutical carrier, eliciting an immune response
CC in a subject, inhibiting the growth of a breast cancer or a prostate
CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to
CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
CC detecting T cells expressing CD8 that specifically recognize the TARP
CC polypeptide in a subject. Eliciting an immune response in a subject
CC comprises administering the polypeptide or polynucleotide, thus producing
CC an immune response in the subject. The immune response comprises a T cell
CC response or inducing cytotoxic T cells that induce lysis of cells
CC expressing the TARP polypeptide. The subject has breast cancer or
CC prostate cancer. The immune response decreases the growth of the prostate
CC cancer or breast cancer. The method further comprises administering an
CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
CC or CTL precursor cells with the polypeptide and an antigen presenting
CC cell to produce activated CTLs matured from the CTL precursors that
CC recognize the breast cancer or the prostate cancer cells, and contacting
CC the breast cancer or the prostate cancer cell with the activated CTLs or
CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
CC breast cancer or the prostate cancer cell. The TARP polypeptide and
CC polynucleotide are useful for treating a subject having breast or
CC prostate cancer. The polypeptide is useful for generating an immune
CC response to breast cancer and prostate cancer cells that express TARP
CC polypeptides. This sequence represents a human TARP polypeptide epitope
CC used in the scope of the invention.
CC
CC
SQ Sequence 9 AA;

Query Match 15.5%; Score 9; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 FLRNFSLML 37
|||||

Db 1 FLRNFSLML 9

RESULT 3
ADM44406
ID ADM44406 standard; peptide; 9 AA.

AC ADM44406;

DT 24-MAR-2005 (first entry)

DE Human TARP polypeptide epitope TARP-22-30.

XX T-cell receptor gamma alternate reading frame protein; TARP;
KW immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;
KW receptor.
XX

OS Homo sapiens.

FN WO200500889-A1.

PD 06-JAN-2005.

PF 02-JUN-2004; 2004WO-US017574.

PR 05-JUN-2003; 2003US-0476467P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Berzofsky JA, Oh S, Pastan I;

DR WPI; 2005-091494/10.

XX New T-cell receptor gamma alternate reading frame protein (TARP)
PT polypeptide, useful for treating a subject having breast or prostate
PT cancer and for generating an immune response to TARP-expressing breast
PT and prostate cancer cells.
XX

PS Example 2; SEQ ID NO 8; 83bp; English.
 XX The invention relates to an immunogenic T-cell receptor gamma alternate
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding
 CC it. The invention also relates to a vector comprising the TARP
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical
 CC composition comprising a therapeutic amount of the polypeptide or the
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response
 CC in a subject, inhibiting the growth of a breast cancer or a prostate
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
 CC detecting T cells expressing CD8 that specifically recognize the TARP
 CC polypeptide in a subject. Eliciting an immune response in a subject
 CC comprises administering the polypeptide or polynucleotide, thus producing
 CC an immune response in the subject. The immune response comprises a T cell
 CC response or inducing cytotoxic T cells that induce lysis of cells
 CC expressing the TARP polypeptide. The subject has breast cancer or
 CC prostate cancer. The immune response decreases the growth of the prostate
 CC cancer or breast cancer. The method further comprises administering an
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
 CC or CTL precursor cells with the polypeptide and an antigen presenting
 CC cell to produce activated CTLs matured from the CTL precursors that
 CC recognize the breast cancer or the prostate cancer cells, and contacting
 CC the breast cancer or the prostate cancer cell with the activated CTLs or
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and
 CC polynucleotide are useful for treating a subject having breast or
 CC prostate cancer. The polypeptide is useful for generating an immune
 CC response to breast cancer and prostate cancer cells that express TARP
 CC polypeptides. This sequence represents a human TARP polypeptide epitope
 CC used in the scope of the invention.
 XX

SQ Sequence 9 AA;

Query Match 15.5%; Score 9; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 RLHETPVFL 30

Db 1 RLHETPVFL 9

RESULT 4
 ADM44405
 ID ADM44405 standard; peptide; 8 AA.
 XX
 AC ADM44405;
 XX
 XX 24-MAR-2005 (first entry)

XX Human TARP polypeptide epitope TARP-2-9.

XX T-cell receptor gamma alternate reading frame protein; TARP;
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;
 KM receptor.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W02005000889-A1.
 PN
 XX 06-JAN-2005.
 PD
 XX 02-JUN-2004; 2004WO-US017574.
 PF
 XX 05-JUN-2003; 2003US-0476467P.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Bezofeky JA, Oh S, Pastan I;
 PI
 XX WPI, 2005-091494/10.

XX New T-cell receptor gamma alternate reading frame protein (TARP)
 PT polypeptide, useful for treating a subject having breast or prostate
 PT cancer and for generating an immune response to TARP-expressing breast
 PT and prostate cancer cells.
 XX
 XX Example 2; SEQ ID NO 7; 83bp; English.
 XX

XX The invention relates to an immunogenic T-cell receptor gamma alternate
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding
 CC it. The invention also relates to a vector comprising the TARP
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical
 CC composition comprising a therapeutic amount of the polypeptide or the
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response
 CC in a subject, inhibiting the growth of a breast cancer or a prostate
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
 CC detecting T cells expressing CD8 that specifically recognize the TARP
 CC polypeptide in a subject. Eliciting an immune response in a subject
 CC comprises administering the polypeptide or polynucleotide, thus producing
 CC an immune response in the subject. The immune response comprises a T cell
 CC response or inducing cytotoxic T cells that induce lysis of cells
 CC expressing the TARP polypeptide. The subject has breast cancer or
 CC prostate cancer. The immune response decreases the growth of the prostate
 CC cancer or breast cancer. The method further comprises administering an
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
 CC or CTL precursor cells with the polypeptide and an antigen presenting
 CC cell to produce activated CTLs matured from the CTL precursors that
 CC recognize the breast cancer or the prostate cancer cells, and contacting
 CC the breast cancer or the prostate cancer cell with the activated CTLs or
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and
 CC polynucleotide are useful for treating a subject having breast or
 CC prostate cancer. The polypeptide is useful for generating an immune
 CC response to breast cancer and prostate cancer cells that express TARP
 CC polypeptides. This sequence represents a human TARP polypeptide epitope
 CC used in the scope of the invention.
 XX

SQ Sequence 8 AA;

Query Match 13.8%; Score 8; DB 9; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QMFPPSPFL 9

Db 1 QMFPPSPFL 8

RESULT 5
 ADM44404
 ID ADM44404 standard; peptide; 9 AA.
 XX
 AC ADM44404;
 XX
 XX 24-MAR-2005 (first entry)

XX Human TARP polypeptide epitope TARP-29-37-9V.

XX T-cell receptor gamma alternate reading frame protein; TARP;
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;
 KM receptor; mutain.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Synthetic.
 OS
 XX

XX Key Location/Qualifiers

FT Misc-difference 9 /note= "Wild-type Leu substituted by Val"

FT
 XX
 PN W02005000889-A1.

PD 06-JAN-2005.
 PF 02-JUN-2004; 2004WO-US017574.
 XX
 PR 05-JUN-2003; 2003US-0476467P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Berzofsky JA, Oh S, Pastan I;
 XX
 DR WPI; 2005-091494/10.
 XX
 PT New T-cell receptor gamma alternate reading frame protein (TARP)
 PT polypeptide, useful for treating a subject having breast or prostate
 PT cancer and for generating an immune response to TARP-expressing breast
 PT and prostate cancer cells.
 XX
 PS
 XX
 Claim 2; SEQ ID NO 6; 83pp; English.
 XX
 CC The invention relates to an immunogenic T-cell receptor gamma alternate
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding
 CC it. The invention also relates to a vector comprising the TARP
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical
 CC composition comprising a therapeutic amount of the polypeptide or the
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response
 CC in a subject, inhibiting the growth of a breast cancer or a prostate
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
 CC detecting T cells expressing CD8 that specifically recognize the TARP
 CC polypeptide in a subject. Eliciting an immune response in a subject
 CC comprises administering the polypeptide or polynucleotide, thus producing
 CC an immune response in the subject. The immune response comprises a T cell
 CC response or inducing cytotoxic T cells that induce lysis of cells
 CC expressing the TARP polypeptide. The subject has breast cancer or
 CC prostate cancer. The immune response decreases the growth of the prostate
 CC cancer or breast cancer. The method further comprises administering an
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
 CC or CTL precursor cells with the polypeptide and an antigen presenting
 CC cell to produce activated CTLs matured from the CTL precursors that
 CC recognize the breast cancer or the prostate cancer cells, and contacting
 CC the breast cancer or the prostate cancer cell with the activated CTLs or
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and
 CC polynucleotide are useful for treating a subject having breast or
 CC prostate cancer. The polypeptide is useful for generating an immune
 CC response to breast cancer and prostate cancer cells that express TARP
 CC polypeptides. This sequence represents a human TARP polypeptide epitope
 CC used in the scope of the invention.
 CC
 XX
 SQ Sequence 9 AA;
 XX
 Query Match 13.8%; Score 8; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 OY 29 FLRNFSLM 36
 DB 1 FLRNFSLM 8
 RESULT 6
 ADH89809
 ID ADH89809 standard; peptide; 9 AA.
 AC ADH89809;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Cell penetrating peptide (CPP) identification method-related peptide 111.
 XX
 KM cell-penetrating peptide; CPP; bulk property value Z-E; Z-E1; Z-E2; Z-E3;
 KM Z-E4; Z-E5; antidiabetic; neuroprotective; nootropic; antiparkinsonian;

XX cardiant; cytoskeletal; tranquilliser; immunosuppressive; antidepressant;
 KW anticouvalant; antiinflammatory; analgesic; neuroleptic;
 KW ophthalmological; antitumor; cell-penetrating; infectious disease;
 KW diabetes type I; diabetes type II; Alzheimer's disease;
 KW Parkinson's disease; cancer; prion disease; cardiovascular disease;
 KW signal transduction.
 XX
 XX
 OS Unidentified.
 XX
 FN WO2003106491-A2.
 XX
 PD 24-DEC-2003.
 XX
 PF 18-JUN-2003; 2003WO-IB003163.
 XX
 PR 18-JUN-2002; 2002SF-00001863.
 XX
 PR 25-JUN-2002; 2002US-0391788P.
 XX
 PA (CEPE-) CEPEP AB.
 XX
 PI Haslbrink M, Pooga M, Metsis M, Kogeran P, Valkna A, Melkas A;
 PI Lindgren M, Gresslund A, Eriksson G, Oesterlsson CG, Budhina M;
 PI Zorko M, Elmqvist A, Soomets U, Lundberg P, Jaarver P, Saar K;
 PI El-Andalousi S, Kilk K, Langel U;
 DR WPI: 2004-090832/09.
 XX
 PT Predicting, designing, detecting, and/or verifying novel cell-penetrating
 PT peptide based on assessment of bulk property value of sequences of cell-
 PT penetrating peptide.
 XX
 PS Disclosure; Page 31, 148pp; English.
 CC This invention relates to a novel method of identifying, designing,
 CC detecting, and/or verifying novel cell-penetrating peptide (CP) based on
 CC assessment of bulk property value Z-B of sequences of CPP comprising 5 or
 CC more individual average interval values Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5,
 CC where Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5 are average values of the
 CC respective descriptor values for the residues in the amino acid sequence.
 CC The invention may be useful for the development of compounds with an
 CC antidiabetic, neuroprotective, nootropic, antiparkinsonian, cardiant,
 CC cytoskeletal, tranquilliser, immunosuppressive, antidepressant,
 CC anticouvalant, antiinflammatory, analgesic, neuroleptic,
 CC ophthalmological or antitumor activity as a stimulator of cell-
 CC penetration. The method of the invention is useful for identifying a cell-
 CC penetrating peptide or protein and/or a cell-penetrating fragment of a
 CC peptide or protein. In addition, the invention may be useful for checking
 CC cellular penetration properties of a peptide, for producing a cell-
 CC penetrating and functional protein-mimicking peptide and for de novo
 CC design and production of an artificial cell-penetrating and/or and
 CC artificial cell-penetrating and functional protein-mimicking peptide.
 CC Compositions developed within the scope of the present invention may be
 CC useful for treating infectious diseases, diabetes type I, diabetes type
 CC II, Alzheimer's disease, Parkinson's disease, cancer, prion disease,
 CC cardiovascular disease or disorders resulting from perturbed signal
 CC transduction. The method of the invention is fast, efficient and reliable
 CC for identifying, detecting, designing CPPs and for screening cellular
 CC uptake of a broad variety of CPPs in vitro and in vivo. The present
 CC sequence is that of a peptide which is related to the invention.
 XX
 XX
 SQ Sequence 9 AA:
 Query Match 10.3%; Score 6; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 FLQLK 17
 Db 4 FLQLK 9
 |||||
 |||||

ID ADM44403 standard; peptide: 9 AA.
 AC ADM44403;
 XX
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Human TARP polypeptide epitope TARP-29-37-3A.
 XX
 XX T-cell receptor gamma alternate reading frame protein; TARP;
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;
 KM receptor; mutuin.
 XX
 XX Homo sapiens.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 3 /note="Wild-type Arg substituted by Ala"
 FT
 XX
 PN MO2005000889-A1.
 XX
 PD 06-JAN-2005.
 XX
 PF 02-JUN-2004; 2004MO-US017574.
 XX
 PR 05-JUN-2003; 2003US-0476467P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Berzofsky JA, Oh S, Pastan I;
 XX WPI; 2005-091494/10.
 DR
 XX
 PT New T-cell receptor gamma alternate reading frame protein (TARP)
 PT polypeptide useful for treating a subject having breast or prostate
 PT cancer and for generating an immune response to TARP-expressing breast
 PT and prostate cancer cells.
 PT
 PS Claim 2; SEQ ID NO 5; 83pp; English.
 XX
 XX The invention relates to an immunogenic T-cell receptor gamma alternate
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding
 CC it. The invention also relates to a vector comprising the TARP
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical
 CC composition comprising a therapeutic amount of the polypeptide or the
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response
 CC in a subject, inhibiting the growth of a breast cancer or a prostate
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
 CC detecting T cells expressing CD8 that specifically recognize the TARP
 CC polypeptide in a subject. Eliciting an immune response in a subject
 CC comprises administering the polypeptide or polynucleotide, thus producing
 CC an immune response in the subject. The immune response comprises a T cell
 CC response or inducing cytotoxic T cells that induce lysis of cells
 CC expressing the TARP polypeptide. The subject has breast cancer or
 CC prostate cancer. The immune response decreases the growth of the prostate
 CC cancer or breast cancer. The method further comprises administering an
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
 CC or CTL precursor cells with the polypeptide and an antigen presenting
 CC cell to produce activated CTLs matured from the CTL precursors that
 CC recognize the breast cancer or the prostate cancer cells, and contacting
 CC the breast cancer or the prostate cancer cell with the activated CTLs or
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and
 CC polynucleotide are useful for treating a subject having breast or
 CC prostate cancer. The polypeptide is useful for generating an immune
 CC response to breast cancer and prostate cancer cells that express TARP
 CC polypeptides. This sequence represents a human TARP polypeptide epitope
 CC used in the scope of the invention.
 CC
 XX Sequence 9 AA;
 SQ

Query Match 10.3%; Score 6; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 NFSJML 37
 DB 4 NFSJML 9

RESULT 8
 AAB06282
 ID AAB06282 standard; peptide: 8 AA.
 AC AAB06282;
 XX
 XX
 DT 03-OCT-2000 (first entry)
 XX
 DE proSKI-1 putative zymogen cleavage site.
 XX
 XX Human; pro-brain-derived neurotrophic factor; proBDNF;
 KM subtilisin-kexin isoenzyme 1; SKI-1; proSKI-1 zymogen processing;
 KM antilipemic; cytostatic; vasotrophic; SKI-1 inhibitor;
 KM hypercholesterolaemia; liver steatosis; Ras-dependent cancer; restenosis;
 KM amyloid protein formation.
 XX
 XX Homo sapiens.
 OS Mus sp.
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site 4..5
 FT
 XX
 PN MO200026348-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 04-NOV-1999; 99MO-CA001058.
 XX
 PR 04-NOV-1998; 98CA-02249648.
 XX
 PA (RECL-) INST RECH CLINIQUES MONTREAL.
 XX
 PI Seidah N, Chretien M, Marcinkiewicz M, Laaksonen R, Davignon J;
 XX WPI; 2000-365601/31.
 DR
 XX
 PT Novel soluble proteic fragment of subtilisin-kexin isoenzyme for
 PT producing a polypeptide useful for treating hypercholesterolemia, liver
 PT steatosis and amyloidosis, comprises a specific amino acid sequence.
 PT
 XX
 PS Example 1; Page 25; 119pp; English.
 XX
 XX The present sequence is a possible cleavage site in proSKI-1 which may be
 CC targeted by a zymogen to form mature subtilisin-kexin isoenzyme 1 (SKI-
 CC 1). The conserved site was identified by alignment of SKI-1 with other
 CC subtilases. SKI-1 is a type-1 membrane-bound proteinase which cleaves pro
 CC -brain-derived neurotrophic factor. Peptides which bind to and are
 CC cleaved by SKI-1 may be used for monitoring SKI-1 activity, for screening
 CC inhibitors of SKI-1 activity, or for screening enhancers of SKI-1
 CC activity. Proteic fragments of SKI-1 which bind to the SKI-1 catalytic
 CC site may be used as inhibitors of SKI-1 activity. They may be used to
 CC treat diseases involving overexpression of SKI-1 or SKI-1 substrate. Such
 CC diseases include hypercholesterolaemia, high levels of fatty acids, liver
 CC lipids or fatty acid pyrophosphate, liver steatosis, Ras-dependent cancer,
 CC restenosis and amyloid protein formation
 CC
 XX Sequence 8 AA;
 SQ
 Query Match 8.6%; Score 5; DB 3; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 SSRLL 23

Db 3 SSRRL 7

RESULT 9
AEA41508
ID AEA41508 standard; peptide; 8 AA.
XX
AC AEA41508;
XX
DT 25-AUG-2005 (first entry)
XX
DE GAG-binding peptide from FAS binding protein #2.
XX
KW Glycosaminoglycan; heparan sulfate; protein engineering; chemokine;
KW inflammation; rheumatoid arthritis; psoriasis; osteoarthritis; asthma;
KW Alzheimer's disease; multiple sclerosis; Antiinflammatory; Antiarthritic;
KW Antirheumatic; Antipsoriatic; Osteopathic; Antiasthmatic;
KW Neutroprotective; Nootropic.
XX
OS Unidentified.
XX
PN WO2005054265-A1.
XX
PD 16-JUN-2005.
XX
PF 02-DEC-2004; 2004WO-EP013670.
XX
PR 04-DEC-2003; 2003AT-00001952.
XX
PA (KUNGL) KUNGL A J.
XX
PI Kungl AJ;
XX
DR WPI, 2005-435363/44.
XX
PT Introducing a glycosaminoglycan (GAG) binding site into a protein, useful
PT for producing mutant chemokines which disrupt chemokine signaling, and
PT can be used to treat inflammatory diseases.
XX
XX Example 6; Page 38; 69pp; English.
XX
CC The invention relates to introducing a glycosaminoglycan (GAG) binding
CC site into a protein comprising identifying a region in the protein which
CC is not essential for structure maintenance, and introducing at least one
CC basic amino acid and/or deleting at least one bulky and/or acidic amino
CC acid in the identified region. The GAG binding site has a GAG binding
CC affinity of Kd less than or equal to 10 microm, preferably less than or
CC equal to 1 microm, still more preferably less than or equal to 0.1
CC microm. Also included are a protein obtainable by the method, a modified
CC GAG binding protein (A) (characterized in that a GAG binding region in
CC the protein is modified by substitution, insertion, and/or deletion of at
CC least one amino acid in order to increase the relative amount of basic
CC amino acids in the GAG binding region, and/or reduce the amount of bulky
CC and/or acidic amino acids in the GAG binding region, preferably at a
CC solvent exposed position, and in that the GAG binding affinity of the
CC protein is increased compared to the GAG binding affinity of a respective
CC wild-type protein), an isolated polynucleic acid molecule encoding the
CC modified protein, a vector comprising an isolated polynucleic acid
CC molecule, a recombinant cell stably transfected with the vector and a
CC pharmaceutical composition comprising a carrier and the protein,
CC polynucleic acid molecule or vector. The protein is a chemokine,
CC preferably IL-8, RANTES or MCP-1. The protein has increased binding
CC affinity to heparan sulfate and/or heparin. A further biologically active
CC region (e.g. for leukocyte activation) is modified, inhibiting or down-
CC regulating the further biological activity of the protein. The protein,
CC polynucleic acid, or vector is useful for inhibiting or suppressing the
CC biological activity of the respective wild-type protein. The protein,
CC polynucleic acid, or vector is useful for producing a medicament for
CC treating an inflammatory condition, preferably a condition selected from
CC rheumatoid arthritis, psoriasis, osteoarthritis, asthma, Alzheimer's
CC disease, and multiple sclerosis. The method provides a means of blocking
CC the interaction between wild-type chemokines and their glycosaminoglycan

CC (GAG) co-receptors, and represents a new strategy for blocking chemokine
CC signaling. The present sequence is a GAG-binding site.
XX
SQ Sequence 8 AA;
XX
Query Match 8.6%; Score 5; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 KKRRA 47
Db 2 KKRRA 6
XX
RESULT 10
AAR59128
ID AAR59128 standard; peptide; 9 AA.
XX
AC AAR59128;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAY-1995 (first entry)
XX
DE Peptide fragment (1.0882) of HBV binds HLA-A2.1.
XX
KW antigen; epitope; immunogenic target protein; PSA; HBVC; HBVs; EBV; HIV1;
KW core antigen; surface antigen; pharmaceutical composition; in vivo;
KW ex vivo; therapeutic; diagnostic; MHC class I molecule;
KW major histocompatibility complex; HLA-A2.1; 9mer; 10mer; anchor;
KW human leukocyte antigen.
XX
OS Hepatitis B virus; strain adw.
XX
PN WO9420127-A1.
XX
PD 15-SEP-1994.
XX
PF 04-MAR-1994; 94WO-US002353.
XX
PR 05-MAR-1993; 93US-00027146.
PR 04-JUN-1993; 93US-00073205.
PR 29-NOV-1993; 93US-00159184.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Grey HM, Sette A, Sidney J, Kast W;
XX
DR WPI; 1994-302678/37.
XX
PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
PT treatment or prophylaxis of cancer, virus infection or autoimmune
PT diseases.
XX
PS Example 5; Page 101; 138pp; English.
XX
CC AAR59104-264 are immunogenic 9mer peptides that contain a HLA-A2.1
CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity
CC of at least 1% as compared to a reference peptide (AAR71293). AAR59128
CC has an IC50 of 0.046 and the sequence occurs at position 1088 in the HBV
CC POL protein. The peptides of the invention can induce cytotoxic T
CC lymphocytes which can react with target cells. They can be used for the
CC treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.
CC (Updated on 25-MAR-2003 to correct PW field.) (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 9 AA;
XX
Query Match 8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 SIMDL 38

Db |||||
5 SLMLL 9

RESULT 11

AAV38039 ID AAV38039 standard; peptide; 9 AA.

AC AAV38039;

DT 29-SEP-1999 (first entry)

XX Hepatitis B virus-derived HLA-binding peptide.

XX Immunogen; HLA; human leukocyte antigen; binding motif; antiviral; MHC;
KM major histocompatibility complex; viral infection; anticancer;
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX Hepatitis B virus.

OS WO9403205-A1.

PN 17-FEB-1994.

PD 06-AUG-1993; 93WO-US007421.

PF 07-AUG-1992; 92US-00926666.

PR 05-MAR-1993; 93US-00027746.

PS (CYTE-) CYTEL CORP.

PI Kubo RT, Grey HM, Sette A, Celis E;

DR WPI; 1994-065403/08.

XX Peptide which specifically binds selected MHC allele - used to induce an
PT immune response for treatment or prevention of viral infection or cancer,
PT or for diagnosis.

PS Disclosure; Page 106; 150pp; English.

XX The sequence is a specific example of a group of new immunogenic peptides
CC having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding motif. For
CC example, the peptides having an HLA-A3.2 binding motif each have 9-10
CC residues and contain, from the N-terminus to the C-terminus, (a) a first
CC conserved residue selected from L, M, I, V, S, A, T, F, C, G, D and B and
CC (b) a second conserved residue from K, R, Y, H or F, where the first and
CC second conserved residues are separated by 6-7 residues. The peptides are
CC capable of binding selected MHC molecules and inducing an immune
CC response. They can be used to treat and/or prevent viral infection and
CC cancer, e.g. prostate cancer, lymphoma, hepatitis or AIDS. They can also
CC be used to produce antibodies for use as diagnostic or therapeutic
CC agents. The peptides can also be used as diagnostic agents

XX Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0;

OY 34 SLMLL 38
1 SLMLL 5

DB

RESULT 12

AAV38106 ID AAV38106 standard; peptide; 9 AA.

AC AAV38106;

DT 29-SEP-1999 (first entry)

DE Hepatitis B virus-derived HLA-binding peptide.

XX Immunogen; HLA; human leukocyte antigen; binding motif; antiviral; MHC;
KM major histocompatibility complex; viral infection; anticancer;
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX Hepatitis B virus.

OS WO9403205-A1.

PN 17-FEB-1994.

PD 06-AUG-1993; 93WO-US007421.

PF 07-AUG-1992; 92US-00926666.

PR 05-MAR-1993; 93US-00027746.

PS (CYTE-) CYTEL CORP.

PI Kubo RT, Grey HM, Sette A, Celis E;

DR WPI; 1994-065403/08.

XX Peptide which specifically binds selected MHC allele - used to induce an
PT immune response for treatment or prevention of viral infection or cancer,
PT or for diagnosis.

PS Disclosure; Page 107; 150pp; English.

XX The sequence is a specific example of a group of new immunogenic peptides
CC having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding motif. For
CC example, the peptides having an HLA-A3.2 binding motif each have 9-10
CC residues and contain, from the N-terminus to the C-terminus, (a) a first
CC conserved residue selected from L, M, I, V, S, A, T, F, C, G, D and B and
CC (b) a second conserved residue from K, R, Y, H or F, where the first and
CC second conserved residues are separated by 6-7 residues. The peptides are
CC capable of binding selected MHC molecules and inducing an immune
CC response. They can be used to treat and/or prevent viral infection and
CC cancer, e.g. prostate cancer, lymphoma, hepatitis or AIDS. They can also
CC be used to produce antibodies for use as diagnostic or therapeutic
CC agents. The peptides can also be used as diagnostic agents

XX Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0;

OY 34 SLMLL 38
3 SLMLL 7

DB

RESULT 13

AAW44612 ID AAW44612 standard; peptide; 9 AA.

AC AAW44612;

DT 27-APR-1998 (first entry)

DE Anti-fungal peptide #213 based on BPI protein (residues 142-169).

XX Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;

XX polymorphonuclear leukocyte; fungicide.

OS Synthetic.

OS Mammalia.

XX Key Modified-site 9 Location/Qualifiers

FT Modified-site 9 /note="C-terminal amide"

XX

PN W09704008-A1.
 XX
 PD 06-FEB-1997.
 XX
 PF 21-MAR-1996; 96WO-US003845.
 XX
 PR 20-JUL-1995; 95US-00504841.
 XX
 PA (XOMA) XOMA CORP.
 XX
 PI Little RG, Lim E, Fadem MB;
 XX
 DR WPI; 1997-132578/12.
 XX
 PR Anti-fungal peptide(s) derived from or based on domain III of
 PT bactericidal/permeability-increasing protein - are used in vitro or in
 CC vivo as a fungicides.
 XX
 PS Claim 1; -pp; 230pp; English.
 XX
 CC This is a specifically claimed anti-fungal peptide which is based on
 CC domain III (amino acids 142-160) of bactericidal-permeability-increasing
 CC protein (BPI), isolated from the granules of mammalian polymorphonuclear
 CC leukocytes. It is used in compositions with diluents, carriers or
 CC adjuvants to treat fungal infections in patients. It may also be used in
 CC vitro to kill or inhibit the replication of fungi, such as in
 CC decontaminating fluids and sterilising medical and implant devices
 XX
 SQ Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17
 |||||
 Db 4 LQLLK 8

RESULT 14
 AAM43777
 ID AAM43777 standard; peptide; 9 AA.
 XX

AC AAM43777;

DT 20-APR-1998 (first entry)

XX Bactericidal/permeability increasing peptide XMP.382.

XX Bactericidal/permeability increasing peptide; BPI; fusion protein;

KW bacterial infection; fungal infection; endotoxin; heparin; angiogenesis;

XX fungicidal; recombinant DNA; vector.

OS Homo sapiens.

OS Synthetic.

XX Key

FT Modified-site

FT 9 Location/Qualifiers

XX W09735009-A1.

PD 25-SEP-1997.

PF 18-MAR-1997; 97WO-US005287.

XX 22-MAR-1996; 96US-00621803.

XX (XOMA) XOMA CORP.

XX Better MD;

DR WPI; 1997-480215/44.

XX
 PT Recombinant production of bactericidal/permeability increasing protein -
 PT by expression as a fusion protein in microbial host cells, then cleaving
 PT the BPI peptide from the carrier.
 XX
 PS Claim 10; Page 136; 186pp; English.

XX
 CC A new recombinant DNA vector construct has been developed which encodes a
 CC fusion protein and is suitable for introduction into a bacterial host.
 CC The vector comprises: (a) DNA encoding at least one cationic
 CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a
 CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site
 CC located between (a) and (b). The present sequence represents a
 CC specifically claimed BPI peptide. The peptides have many uses including
 CC the treatment of bacterial and fungal infections. BPI peptides also bind
 CC to endotoxins and heparin, neutralising their effects. The peptides have
 CC further been shown to inhibit angiogenesis (partly due to heparin-binding
 CC activity). The fusion proteins have been found to be expressed in large
 CC amounts without significant proteolysis, and in some cases are actually
 CC secreted from the host cells. This allows the indirect production of anti-
 CC microbial BPI peptides in microbial hosts

SQ Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17
 |||||
 Db 4 LQLLK 8

RESULT 15
 AAY45608
 ID AAY45608 standard; peptide; 9 AA.
 XX

AC AAY45608;

DT 01-DEC-1999 (first entry)

XX Immunogenic peptide having a human leukocyte antigen binding motif #219.

XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;

KW immune response; T cell activation; major histocompatibility complex;

KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;

KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;

KW vaccine; immunisation.

XX Synthetic.

OS Homo sapiens.

PN W09945954-A1.

PD 16-SEP-1999.

PF 13-MAR-1998; 98WO-US005039.

XX 13-MAR-1998; 98WO-US005039.

XX (EPI-M-) EPIMUNE INC.

PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX WPI; 1999-551214/46.

XX New immunogenic peptides with HLA binding motif, useful in treatment and

XX diagnosis of cancers and viral diseases.

XX Claim 1; Page 37; 150pp; English.

XX AAY45390 to AAY48214 represent specifically claimed immunogenic peptides

CC having a human major histocompatibility complex (MHC) Class I (also known

CC as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides
 CC can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2
 CC or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against
 CC the antigen from which the peptide is derived. Cytotoxic T lymphocytes
 CC (CTLs) which destroy antigen-bearing cells are normally induced by an
 CC antigen in the form of a peptide fragment bound to a HLA molecule, rather
 CC than the intact foreign antigen itself, and are particularly important in
 CC tumour rejection and in fighting viral infections. The peptides are
 CC therefore useful therapeutically to treat or prevent viral infections and
 CC cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B
 CC and C, AIDS, and renal carcinoma. They can be administered as vaccines to
 CC elicit an immune response in individuals susceptible or otherwise at risk
 CC of viral infection or cancer, or used to treat chronic or acute
 CC conditions. They are also useful diagnostically, and can be used to
 CC induce a cytotoxic T cell response, by contacting a cytotoxic T cell with
 CC the peptide e.g. to produce CTLs ex vivo for infusion back into a
 CC patient. The polynucleotides encoding the immunogenic peptides are also
 CC useful therapeutically and for immunisation as above
 CC
 XX

SQ Sequence 9 AA;

Query Match

Best Local Similarity 8.6%; Score 5; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SLML 38

DB 1 SLML 5

Search completed: December 13, 2005, 08:21:44
 Job time : 183 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:36 ; Search time 37 Seconds
(without alignments)
150.826 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58

Sequence: 1 MQMPPSPPLFFFLQLKQSS.....RYIGKRRRATRFMDPRRGTTP 58

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size: 0

Total number of hits satisfying chosen parameters: 645

Minimum DB seq length: 8
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database:

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	6.9	8	2	PC4131
2	4	6.9	10	2	A56633
3	4	6.9	10	2	A32543
4	3	5.2	8	2	A44960
5	3	5.2	8	2	A61348
6	3	5.2	8	2	S08995
7	3	5.2	8	2	A49823
8	3	5.2	8	2	A28004
9	3	5.2	8	2	A43976
10	3	5.2	8	2	B43976
11	3	5.2	8	2	T10077
12	3	5.2	8	2	S19288
13	3	5.2	8	2	S21288
14	3	5.2	8	2	A05169
15	3	5.2	8	2	JS0316
16	3	5.2	8	2	S71919
17	3	5.2	8	2	A61102
18	3	5.2	9	2	PT0326
19	3	5.2	9	2	S19523
20	3	5.2	9	2	A31576
21	3	5.2	9	2	PC7076
22	3	5.2	10	2	A60624
23	3	5.2	10	2	A60410
24	3	5.2	10	2	S08997
25	3	5.2	10	2	A60421
26	3	5.2	10	2	S08998
27	3	5.2	10	2	A26381
28	3	5.2	10	2	JN0440
29	3	5.2	10	2	PQ0753

30	3	5.2	10	2	A61622	vitellogenin, 190k
31	3	5.2	10	2	C39111	Ig heavy chain C r
32	3	5.2	10	2	S10785	enamelin, 22K - bo
33	3	5.2	10	2	S70722	65.4K GTP-binding
34	3	5.2	10	2	PT0084	protein QA600021 -
35	3	5.2	10	2	B59272	peptide-N4-(N-acet
36	3	5.2	10	2	A43977	transcription fact
37	3	5.2	10	2	A42089	angiotensin precu
38	3	5.2	10	2	A90917	cytochrome-c oxida
39	3	5.2	10	2	T17075	neuropeptide Pec-H
40	3	5.2	10	2	S53789	NADH2 dehydrogenas
41	3	5.2	10	2	PQ0784	sperm-activating p
42	3	5.2	10	2	G60787	sperm-activating p
43	3	5.2	10	2	G60787	sperm-activating p
44	3	5.2	10	2	G60787	sperm-activating p
45	3	5.2	10	2	C60588	sperm-activating p

ALIGNMENTS

RESULT 1

PC4131
Hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)

C:Species: Pseudomonas aeruginosa

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 31-Dec-2004

C:Accession: PC4131

R:Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.

Gene 167, 87-91, 1995

A:Title: Sequencing and characterization of the downstream region of the genes encoding

Y for biosynthesis of heme d1.

A:Reference number: J04552; PMID:96144254; PMID:8566817

A:Accession: PC4131

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-8 <RAW>

A:Cross-references: UNIPROT:P95412; UNIPARC:UPI000017A96A; DDBJ:D50473; MID:g1217594

A:Note: this ORF is not annotated in GenBank entry PSENIIRC, release 113.0

C:Superfamily: Pseudomonas stutzeri nird protein

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 8;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 SRRL 23
Db 5 SRRL 8

RESULT 2

A56633
neomysuppressin - flesh fly (Sarcophaga bullata)

N:Alternate names: Neb-MS

C:Species: Sarcophaga bullata

C>Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004

C:Accession: A56633

R:Fonagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; Bueds, H.; De Loof, A.

Comp. Biochem. Physiol. C 102, 239-245, 1992

A:Title: Isolation, primary structure and synthesis of neomysuppressin, a myoinhibiting

A:Reference number: A56633; PMID:93047886; PMID:1358537

A:Accession: A56633

A:Molecule type: protein

A:Residues: 1-10 <FOR>

A:Cross-references: UNIPROT:P61850; UNIPARC:UPI000003AD0A

A:Experimental source: head

A:Note: sequence extracted from NCBI backbone (NCBIP:119072)

C:Keywords: amidated carboxyl end; neuropeptide

F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLLR 31
|||
Db 6 VFLLR 9

RESULT 3

A32543
cardioexcitatory neuropeptide - desert locust
C/Species: Schistocerca gregaria (desert locust)
C/Date: 20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change 05-Oct-2004
C/Accession: A32543
R/Robb, S.; Packman, L.C.; Evans, P.D.
Biochem. Biophys. Res. Commun. 160, 850-856, 1989
A/Title: Isolation, primary structure and bioactivity of Schistocerca gregaria, a FMRF-amide
A/Reference number: A32543; MUID:89246543; PMID:2719702
A/Accession: A32543
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-10 <ROB>
A/Cross-references: UNIPROT:P38553; UNIPARC:UPI000003AD64
C/Keywords: amidated carboxyl end; neuropeptide
F10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 6.9%; Score 4; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLLR 31
|||
Db 6 VFLLR 9

RESULT 4

A44960
neuropeptide Ied-CC-I - Colorado potato beetle
C/Species: Leptinotarsa decemlineata (Colorado potato beetle)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A44960
R/Gaede, G.; Kellner, R.
Peptides 10, 1287-1289, 1989
A/Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and
A/Reference number: A44960; MUID:90160053; PMID:2576128
A/Accession: A44960
A/Molecule type: protein
A/Residues: 1-8 <GAB>
A/Cross-references: UNIPROT:P04548; UNIPARC:UPI000012CDBB

C/Superfamily: adipokinetic hormone
C/Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F11/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F18/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34
|||
Db 3 NFS 5

RESULT 5

A61348
red pigment-concentrating hormone - northern shrimp
N/Alternate names: blanching hormone
C/Species: Penaeus borealis (northern shrimp)
C/Date: 02-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 09-Jul-2004
C/Accession: A61348; S07139
R/Fernlund, P.; Josefsson, L.
Science 177, 173-175, 1972
A/Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.

A/Reference number: A61348; MUID:72228738; PMID:5041363
A/Accession: A61348
A/Molecule type: protein

A/Residues: 1-8 <FER1>
A/Cross-references: UNIPROT:P08939; UNIPARC:UPI000013460C
R/Fernlund, P.
Biochim. Biophys. Acta 371, 304-311, 1974
A/Title: Structure of the red-pigment-concentrating hormone of the shrimp, Penaeus borealis
A/Reference number: S07139; MUID:75054965; PMID:4433569
A/Accession: S07139
A/Molecule type: protein
A/Residues: 'E', 2-8 <FER2>
A/Cross-references: UNIPARC:UPI000017661E

A/Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C/Comment: This neuropeptide, isolated from the eyestalks of the shrimp and active in pig-
mented pigment-containing cells.

C/Superfamily: adipokinetic hormone
C/Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutamic
F11/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F18/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34
|||
Db 3 NFS 5

RESULT 6

S08995
hypertrehalosemic hormone I - oriental cockroach
N/Alternate names: Pea-CAH-I
C/Species: Blatta orientalis (oriental cockroach)
C/Date: 30-Jun-1992 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
C/Accession: S08995
R/Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990

A/Title: Primary structures of hypertrehalosemic neuropeptides isolated from the corpor-
entails and of the stick insect Ectatopsoma tlaxiutum assigned by tandem fast atom bombard

A/Reference number: S08995; MUID:90253659; PMID:2340112
A/Accession: S08995
A/Molecule type: protein
A/Residues: 1-8 <GAB>
A/Cross-references: UNIPROT:P04548; UNIPARC:UPI000012CDBB

A/Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C/Superfamily: adipokinetic hormone
C/Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F11/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F18/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34
|||
Db 3 NFS 5

RESULT 7

A49823
adipokinetic hormone I - American cockroach

N/Alternate names: periplanetin CC-1
C/Species: Periplaneta americana (American cockroach)

C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C/Accession: A49823
R/Scarborough, R.M.; Jamieson, G.C.; Kallish, F.; Kramer, S.J.; McBrice, G.A.; Miller, C.
Proc. Natl. Acad. Sci. U.S.A. 81, 5575-5579, 1984

A/Title: Isolation and primary structure of two peptides with cardioacceleratory and hyp

A/Reference number: A49823; MUID:84298179; PMID:6591205
A/Accession: A49823
A/Molecule type: protein
A/Residues: 1-8 <SCA>
A/Cross-references: UNIPROT:P04548; UNIPARC:UPI000012CDBB

C:Superfamily: adipokinetin hormone
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34
|||
Db 3 NFS 5

RESULT 8
A28004
adipokinetin hormone G - two-spotted cricket

N:Alternate names: AKH-G
C:Species: Gryllus bimaculatus (two-spotted cricket)

C>Date: 30-Jun-1989 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004

C:Accession: A28004

R:Gade, G.; Rinehart, K.L.

Biochem. Biophys. Res. Commun. 149, 908-914, 1987

A:Title: Primary sequence analysis by fast atom bombardment mass spectrometry of a peptide

A:Reference number: A28004; PMID:8810653; PMID:3426616

A:Accession: A28004

A:Molecule type: protein

A:Residues: 1-8 <GAE>

A:Cross-references: UNIPROT:P14086; UNIPARC:UPI000012576F

A>Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have

C:Superfamily: adipokinetin hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34
|||
Db 3 NFS 5

RESULT 9
A43976

hypertrehalosemic hormone - yellow mealworm

C:Species: Tenebrio molitor (yellow mealworm)

C>Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 09-Jul-2004

C:Accession: A43976

R:Gade, G.; Rosinski, G.

Peptides 11, 455-459, 1990

A:Title: The primary structure of the hypertrehalosemic neuropeptide from tenebrionid be

A:Reference number: A43976; PMID:90341081; PMID:2381871

A:Accession: A43976

A:Molecule type: protein

A:Residues: 1-8 <GAE>

A:Cross-references: UNIPROT:P25419; UNIPARC:UPI000012CDC8

C:Superfamily: adipokinetin hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34
|||
Db 3 NFS 5

RESULT 10

B43976
hypertrehalosemic hormone - beetle (Zophobas rugipes)

C:Species: Zophobas rugipes

C>Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 09-Jul-2004

C:Accession: B43976

R:Gade, G.; Rosinski, G.

Peptides 11, 455-459, 1990

A:Title: The primary structure of the hypertrehalosemic neuropeptide from tenebrionid be

A:Reference number: A43976; PMID:90341081; PMID:2381871

A:Accession: B43976

A:Molecule type: protein

A:Residues: 1-8 <GAE>

A:Cross-references: UNIPROT:P25419; UNIPARC:UPI000012CDC8

C:Superfamily: adipokinetin hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34
|||
Db 3 NFS 5

RESULT 11
T10077

hypothetical protein N - Methylophilus methylotrophus (fragment)

C:Species: Methylophilus methylotrophus

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T10077

R:Christoserdov, A.V.; McIntire, W.S.; Mathews, F.S.; Lidstrom, M.E.

J. Bacteriol. 176, 4073-4080, 1994

A:Title: Organization of the methylamine utilization (mau) genes in Methylophilus methyl

A:Reference number: Z16936; PMID:9429242; PMID:8021188

A:Accession: T10077

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-8 <CHI>

A:Cross-references: UNIPARC:UPI000011E8A1; EMBL:L26407; NID:G561931; PIRN:AAB46955.1; PIR

C:Experimental source: strain W3A1

C:Genetics:

A:Gene: mauN

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 RRL 23
|||
Db 2 RRL 4

RESULT 12
S19288

acylase - Kluyvera cryocrescens

C:Species: Kluyvera cryocrescens

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S19288

R:Martin, U.; Slade, A.; Aitken, A.; Arche, R.; Vidén, R.

Biochem. J. 280, 659-662, 1991

A:Title: Chemical modification of serine at the active site of penicillin acylase from KJ

A:Reference number: S19288; PMID:1764029

A:Accession: S19288

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-8 <MAR>

A:Cross-references: UNIPROT:Q7M124; UNIPARC:UPI000017AA77

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 IGX 43
|||
Db 6 IGX 8

RESULT 13

S21288
lectin - potato (fragment)
C:Species: Solanum tuberosum (potato)
C:Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: S21288
R:Miller, D.J.; Allen, A.K.; Smith, C.G.; Sidebottom, C.; Slabas, A.R.; Bolwell, G.P.
Biochem. J. 283, 813-821, 1992
A:Title: Chitin-binding proteins in potato (Solanum tuberosum L.) tuber. Characterization
A:Reference number: S21288; MUID:92272683; PMID:1590771
A:Accession: S21288
A:Molecule type: protein
A:Residues: 1-8 <MIL>
A:Cross-references: UNIPROT:Q7M1V6; UNIPARC:UPI000017B0BF
A:Experimental source: var. Ulster Sceptre
C:Function:
A:Description: may be involved in defence mechanism of the plant
C:Keywords: hydroxyproline; lectin

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PSP 8
|||
Db 5 PSP 7

RESULT 14

A05169
neuropeptide M-I - American cockroach
C:Species: Periplaneta americana (American cockroach)
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C:Accession: A05169
R:Witten, J.L.; Schaffer, M.H.; O'Shea, M.; Cook, J.C.; Hemling, M.E.; Rinehart Jr., K.L.
Biochem. Biophys. Res. Commun. 124, 350-358, 1984
A:Title: Structures of two cockroach neuropeptides assigned by fast atom bombardment mass
A:Reference number: A0118; MUID:85046530; PMID:6548628
A:Accession: A05169
A:Molecule type: protein
A:Residues: 1-8 <WIT>
A:Cross-references: UNIPROT:P04548; UNIPARC:UPI000017BE15
C:Keywords: neuropeptide

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34
|||
Db 3 NFS 5

RESULT 15

JS0316
leucokinin VI - Madeira cockroach
C:Species: Leucophaea maderae (Madeira cockroach)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: JS0316
R:Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol. C 89, 27-30, 1987
A:Title: Isolation, primary structure, and synthesis of leucokinin V and VI: myotropic
A:Reference number: JS0315
A:Accession: JS0316
A:Molecule type: protein
A:Residues: 1-8 <HOL>

A:Cross-references: UNIPROT:P19988; UNIPARC:UPI000012E29E
C:Comment: Leucokinin, a family of cephalomyotropic peptides, stimulate contractile act
C:Keywords: amidated carboxyl end; cephalomyotropic peptide; pyroglutamic acid
F:1/Modified site: pyroglutamic acid (Gln) #status experimental
F:8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 QSS 20
|||
Db 1 QSS 3

Search completed: December 13, 2005, 08:27:10
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: December 13, 2005, 08:18:33 ; Search time 226 Seconds

(Without alignments)
181.065 Million cell updates/sec

Title: US-10-031-158b-14

Perfect score: 58
Sequence: 1 MQMFPSPPLFFFLQLKQSS.....RYIGKKRRATRFMDPRRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 2166443 seqs, 705528306 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2877

Minimum DB seq length: 8
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	6.9	8	015896 BABBO	015896 babesia boy
2	4	6.9	9	088889 MOUSE	088889 mus musculu
3	4	6.9	10	FARP_LOCHI	P84306 locusta mlg
4	4	6.9	10	FARP_SCHGR	P84307 schistocerc
5	4	6.9	10	LCMS_LEBWA	P11144 leucophaea
6	4	6.9	10	NEMS_SARBU	P18850 sarcophaga
7	4	6.9	10	06A3T0_9FILI	06a3t0 archangiopt
8	4	6.9	10	Q76M18_9TELE	Q76m18 euryparynx
9	3	5.2	8	AKHG_GRYBI	P67785 gryllus bim
10	3	5.2	8	AKH_ROMMI	P67786 romalea mic
11	3	5.2	8	CADI_ENTFA	P13268 enterococcu
12	3	5.2	8	FARI_PANRE	P1872 panagrellus
13	3	5.2	8	FARI_PENMO	P83316 penaeus mon
14	3	5.2	8	FAR2_MACRS	P83275 macrodrachi
15	3	5.2	8	FAR3_HOMAM	P41486 homarus ame
16	3	5.2	8	FAR4_HOMAM	P41487 homarus ame
17	3	5.2	8	HTF1_BLAOR	P84261 blattia orie
18	3	5.2	8	HTF1_LEPDE	P84260 lepidinocera
19	3	5.2	8	HTF1_PERAM	P84259 periplaneta
20	3	5.2	8	HTF_TENNO	P87788 tenebrio mo
21	3	5.2	8	HTF_ZOPRU	P67790 zophobas ru
22	3	5.2	8	LC6K_LEBWA	P19988 leucophaea
23	3	5.2	8	RPCH_PANBO	P08939 pandanus bo
24	3	5.2	8	013591_YEAST	013591 saccharomyc
25	3	5.2	8	P82858_9BASI	P82858 puccinia re
26	3	5.2	8	Q69YH8_HUMAN	Q69yh8 homo sapien
27	3	5.2	8	Q8TF70_HUMAN	Q8tf70 homo sapien
28	3	5.2	8	Q75MD1_HUMAN	Q75md1 homo sapien
29	3	5.2	8	Q86B59_STRPU	Q86b59 strongyloce
30	3	5.2	8	Q37854_BPR17	Q37854 bacterioph
31	3	5.2	8	Q40530_TOBAC	Q40530 nicotiana t

32	3	5.2	8	06JC68 SOYEN	06jc68 glycine max
33	3	5.2	8	070Y84_9LAMI	070y84 plectranthu
34	3	5.2	8	07MIV6_SOLMU	07miv6 solanum tub
35	3	5.2	8	Q9RS80_SHIDY	Q9rs80 shigella dy
36	3	5.2	8	Q79F63_SYNSP	Q79f63 synedrococc
37	3	5.2	8	Q7M124_KLUCI	Q7m124 kluyvera ci
38	3	5.2	8	Q78DX6_RAT	Q78dx6 rattus norv
39	3	5.2	8	08R5M9_MOUSE	08r5m9 mus musculu
40	3	5.2	8	Q9ET16_MESAU	Q9et16 mesocricetu
41	3	5.2	8	Q9ET17_MUSCR	Q9et17 mus caroli
42	3	5.2	8	Q9ET18_MUSSP	Q9et18 mus spretus
43	3	5.2	8	Q6PUD5_SV40	Q6pud5 simian viru
44	3	5.2	8	Q6PUD7_SV40	Q6pud7 simian viru
45	3	5.2	8	Q6PUD9_SV40	Q6pud9 simian viru

ALIGNMENTS

RESULT 1
ID 015896 BABBO PRELIMINARY; PRT; 8 AA.
AC 015896;
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-FEB-2005 (TREMREL. 29, Last annotation update)
DE 12D3 antigen (Fragment).
GN Name=12D3;
OS Babesia bovis.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.
OX NCBI_TaxID=5865;
RN [1]
RP NCLEOTIDE SEQUENCE.
RC STRAIN=Samford attenuated; Riddles P.W.;
RA Silline G.U., Blakeley R.L., Riddles P.W.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NCLEOTIDE SEQUENCE.
RA Silline G., Blakeley R., Riddles P.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U4917; BAB6362.1; -; Genomic DNA.
DR EMBL; AY170919; AAN64587.1; -; Genomic DNA.
FT NON TER 8
SQ SEQUENCE 8 AA; 984 MW; FDD2C9D411ADD726 CRC64;
Query Match 6.9%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 ATRF 50
Db 3 ATRF 6
RESULT 2
ID 088889 MOUSE PRELIMINARY; PRT; 9 AA.
AC 088889;
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-MAR-2004 (TREMREL. 26, Last annotation update)
DE Ubiquitin-conjugating enzyme UBCH4 (Fragment).
GN Name=Ube213; Synonyms=UbcM4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NCLEOTIDE SEQUENCE.
RC STRAIN=129/Sv;
RX MEDLINE=7057256; PubMed=8901595; DOI=10.1073/pnas.93.22.12412;
RA Harbers K., Mueller U., Grams A., Li E., Jaenisch R., Franz T.;

RT "Protein integration into a gene encoding a ubiquitin-conjugating
 RT enzyme results in a placental defect and embryonic lethality."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12412-12417(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN-128/SV;
 RA MEDLINE=99132641;
 RM Mller U., Grams A., Martinez-Noel G., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Harbers K.;
 RT "Structure of the gene encoding the ubiquitin-conjugating enzyme
 RT UbcM4, characterization of its promoter, and chromosomal location."
 RL Gene 224:109-116(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN-129/SV;
 RA Mueller U., Grams A., Martinez-Noel G., Harbers K.;
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AF071557; AAD0128.1; -, Genomic_DNA.
 DR MGI; MGI:109240; Ube213.
 FT NON TER
 SQ SEQUENCE 9 AA; 1063 MW; C90F9734415BDD CRC64;
 Query Match 6.9%; Score 4; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.2e+06; Mismatches 0; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 20 SRRL 23
 Db 4 SRRL 7
 RESULT 3
 FARP LOCMI STANDARD; PRT; 10 AA.
 AC P84307; P38553;
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-2005 (Rel. 46, Last annotation update)
 DE Schistocerca gregaria (Schistocerca gregaria) (Cardioexcitatory neuropeptide).
 OS Locusta migratoria (Migratory locust).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 CC Acridoidea; Acrididae; Oedipodinae; Locusta.
 CC NCBI_TaxID=7004;
 RN [1]
 RP PROTEIN SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RC TISSUE=Corpora cardiaca;
 RX MEDLINE=93324430; PubMed=7687352; DOI=10.1016/0196-9781(93)90126-2;
 RA Schoofs L., Holman G.M., Paemen L., Veelaert D., Amelincx M.,
 RA de Loof A.;
 RT "Isolation, identification, and synthesis of PVDHFLRFamide
 RT (Schistocerca gregaria) in Locusta migratoria and its association with the
 RT male accessory glands, the salivary glands, the heart, and the
 RT oviduct."
 RL Peptides 14:409-421(1993).
 CC -1- FUNCTION: Muscle inhibiting agent. Involved in the neural control
 CC of the visceral muscles of the heart, accessory glands and
 CC oviduct. May be involved in the regulation of saliva secretion.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Found in axons of the male accessory glands,
 CC the salivary glands, the heart, and the oviduct.
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
 CC family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC Amidation: Direct protein sequencing; Neuropeptide.
 CC MOD_RES 10 10 Phenylalanine amide.
 FT

SQ SEQUENCE 10 AA; 1244 MW; D3C51729D2C1EAB2 CRC64;
 Query Match 6.9%; Score 4; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.2e+03; Mismatches 0; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 28 VFRL 31
 Db 6 VFRL 9
 RESULT 4
 FARP SCHGR STANDARD; PRT; 10 AA.
 AC P84307; P38553;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-2005 (Rel. 46, Last annotation update)
 DE Schistocerca gregaria (Desert locust).
 OS Schistocerca gregaria (Desert locust).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 CC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
 CC NCBI_TaxID=7010;
 RN [1]
 RP PROTEIN SEQUENCE, AND FUNCTION.
 RC TISSUE=Thoracic nervous system;
 RX MEDLINE=89246543; PubMed=2719702;
 RA Robb S., Packman L.C., Evans P.D.;
 RT "Isolation, primary structure and bioactivity of schistocerca gregaria, a
 RT FMRF-amide-like neuropeptide from the locust, Schistocerca gregaria."
 RL Biochem. Biophys. Res. Commun. 160:850-856(1989).
 CC -1- FUNCTION: Muscle inhibiting agent. Involved in the neural control
 CC of the visceral muscles of the heart, accessory glands and
 CC oviduct. May be involved in the regulation of saliva secretion.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
 CC family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC PIR: A32543; A32543.
 DR Amidation: Direct protein sequencing; Neuropeptide.
 FT MOD_RES 10 10 Phenylalanine amide.
 SQ SEQUENCE 10 AA; 1244 MW; D3C51729D2C1EAB2 CRC64;
 Query Match 6.9%; Score 4; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.2e+03; Mismatches 0; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 28 VFRL 31
 Db 6 VFRL 9
 RESULT 5
 LCMS_LERUA STANDARD; PRT; 10 AA.
 ID LCMS_LERUA
 AC P21144; P41497;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Leucomyosuppressin (LMS) (Lem-MS).
 OS Leucomyosuppressin (LMS) (Lem-MS).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattellidae;
 CC Blaberidae; Leucomyosuppressin.
 CC NCBI_TaxID=6988;
 RN [1]

RP PROTEIN SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure and synthesis of leucenyosuppressin, an insect neuropeptide that inhibits spontaneous contractions of the cockroach hindgut";
 RL Somp. Biochem. Physiol. 85C:329-333(1986).
 CC -1- FUNCTION: Inhibits the spontaneous contractions of cockroach proctoderm (hindgut).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 KW Amidation; Direct protein sequencing; Neuropeptide;
 KM Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 Pyrrolidone carboxylic acid.
 FT MOD_RES 10 10 Phenylalanine amide.
 SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;
 Query Match 6.9%; Score 4; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 28 VFLLR 31
 Db 6 VFLLR 9

RESULT 6
 NEMS_SARBU STANDARD; PRT; 10 AA.
 ID NEMS_SARBU
 AC P61850; P41494; Q9VC91;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Neomysosuppressin (Neb-MS) (TIDVHVFLLRamide).
 GN Name=NEMS;
 OS Sarcophaga bullata (Grey flesh fly) (Neobelliella bullata).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 CC Sarcophagidae; Sarcophaga; Neobelliella.
 CC NCBI_TaxID=7385;
 RN [1]
 RP PROTEIN SEQUENCE.
 RC TISSUE=Head;
 RX MEDLINE=93047886; PubMed=1358537;
 RA Fonagy A., Schoofs L., Proost P., Van Damme J., Bueda H., De Loof A.;
 RT "Isolation, primary structure and synthesis of neomysosuppressin, a myoinhibiting neuropeptide from the grey fleshfly, Neobelliella bullata";
 RL Somp. Biochem. Physiol. 102C:239-245(1992).
 CC -1- FUNCTION: Myoinhibiting neuropeptide.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 KW Amidation; Direct protein sequencing; Neuropeptide.
 FT MOD_RES 10 10 Phenylalanine amide.
 SQ SEQUENCE 10 AA; 1248 MW; D3C00329D2C1EAB2 CRC64;
 Query Match 6.9%; Score 4; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLLR 31
 Db 6 VFLLR 9

RESULT 7
 ID 06A370_9FILLI PRELIMINARY; PRT; 10 AA.
 AC 06A370;
 DT 25-OCT-2004 (TEMBLrel. 28, Created)
 DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
 DE ATP synthase beta subunit (EC 3.6.3.14) (fragment).
 GN Name=atp3;
 OS Archanglopteris somai.
 CC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Monilliformes; Filicophyta; Marattiales; Marattiales;
 CC Marattiaceae; Archanglopteris.
 CC NCBI_TaxID=203826;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Chiang T., Chiang Y., Chou C., Cheng Y., Chiou W.;
 RT "Phylogeny and conservation of Archanglopteris somai and A. itoi (Marattiales, Pteridophyta) based on nucleotide variation of cpDNA atp3-rbcL intergenic spacer";
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Chiang Y.C.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ505259; CAD4048.1; -; Genomic_DNA.
 DR GO; GO:0009507; C:chloroplast; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 KW Chloroplast; Hydrolyase.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1195 MW; 9E2AF0C9C7376451 CRC64;
 Query Match 6.9%; Score 4; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 PLFF 11
 Db 5 PLFF 8

RESULT 8
 ID 076ML8_9TELE PRELIMINARY; PRT; 10 AA.
 AC 076ML8;
 DT 05-JUL-2004 (TEMBLrel. 27, Created)
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TEMBLrel. 29, Last annotation update)
 DE NMDH dehydrogenase subunit 6 (Fragment).
 GN Name=ND6;
 OS Burypharynx pelicanoides (pelican eel).
 CC Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Anguilliformes;
 CC Eurypharyngidae; Burypharynx.
 CC NCBI_TaxID=55117;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22967687; PubMed=12949142; DOI=10.1093/molbev/meg206;
 RA Inoue J.G., Miya M., Tsukamoto K., Nishida M.;
 RT "Evolution of the deep-sea gulper eel mitochondrial genomes: large-scale gene rearrangements originated within the eels";
 RL Mol. Biol. Evol. 20:1917-1924(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Inoue J., Inoue J., Miya M., Tsukamoto K., Nishida M.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; AB046481; BAB87147.1; -; Genomic_DNA.
DR EMBL; AB046485; BAB87155.1; -; Genomic_DNA.
DR EMBL; AB046489; BAB87163.1; -; Genomic_DNA.
DR EMBL; AB046477; BAB87139.1; -; Genomic_DNA.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1297 MW; D43ECB9C9C059C9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 33 FSLM 36
Db 7 FSLM 10

RESULT 9
AKHG_GRYBI STANDARD; PRT; 8 AA.
ID AKHG_GRYBI
AC P67785; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone G (AKH-G).
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OC NCBI_TaxID=6999;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
RT of a peptide with adipokinetic activity from the corpora cardiaca of
RT the cricket Gryllus bimaculatus."
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the AKH/HRTN/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC PIR: A28004; A28004.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroliidone carboxylic acid.
FT MOD_RES 1 1 Pyroliidone carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 10
AKH_ROMMI STANDARD; PRT; 8 AA.
ID AKH_ROMMI
AC P67786; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone (AKH) (RO II).
OS Romalea microptera (lubber grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caellifera; Acridomorpha;
OC Acridoidea; Romaleidae; Romalea.
OC NCBI_TaxID=7007;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948; DOI=10.1016/0196-9781(88)90107-6;
RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH family from
RT the lubber grasshopper, Romalea microptera."
RL Peptides 9:681-688(1988).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted; HRTN/RPCH family.
CC -1- SIMILARITY: Belongs to the AKH/HRTN/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroliidone carboxylic acid.
FT MOD_RES 1 1 Pyroliidone carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 11
CAD1_ENTFA STANDARD; PRT; 8 AA.
ID CAD1_ENTFA
AC P13268;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Sex pheromone CAD1.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OC NCBI_TaxID=151;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=85051889; PubMed=6437872; DOI=10.1016/0014-5793(84)81248-X;
RA Mori M., Sagakami Y., Narita M., Isegai A., Fujino M., Kitada C.,
RA Craig R.A., Clevellel D.B., Suzuki A.;
RT "Isolation and structure of the bacterial sex pheromone, CAD1, that
RT induces plasmid transfer in Streptococcus faecalis."
RL FEBS Lett. 178:97-100(1984).
CC -1- FUNCTION: Involved in the conjugative transfer of the hemolysin
CC plasmid PAD1.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
DR EMBL; AB046481; BAB87147.1; -; Genomic_DNA.
DR EMBL; AB046485; BAB87155.1; -; Genomic_DNA.
DR EMBL; AB046489; BAB87163.1; -; Genomic_DNA.
DR EMBL; AB046477; BAB87139.1; -; Genomic_DNA.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1297 MW; D43ECB9C9C059C9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 33 FSLM 36
Db 7 FSLM 10

RESULT 9
AKHG_GRYBI STANDARD; PRT; 8 AA.
ID AKHG_GRYBI
AC P67785; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone G (AKH-G).
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OC NCBI_TaxID=6999;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
RT of a peptide with adipokinetic activity from the corpora cardiaca of
RT the cricket Gryllus bimaculatus."
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the AKH/HRTN/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC PIR: A28004; A28004.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroliidone carboxylic acid.
FT MOD_RES 1 1 Pyroliidone carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 10
AKH_ROMMI STANDARD; PRT; 8 AA.
ID AKH_ROMMI
AC P67786; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone (AKH) (RO II).
OS Romalea microptera (lubber grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caellifera; Acridomorpha;
OC Acridoidea; Romaleidae; Romalea.
OC NCBI_TaxID=7007;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948; DOI=10.1016/0196-9781(88)90107-6;
RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH family from
RT the lubber grasshopper, Romalea microptera."
RL Peptides 9:681-688(1988).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted; HRTN/RPCH family.
CC -1- SIMILARITY: Belongs to the AKH/HRTN/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroliidone carboxylic acid.
FT MOD_RES 1 1 Pyroliidone carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 11
CAD1_ENTFA STANDARD; PRT; 8 AA.
ID CAD1_ENTFA
AC P13268;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Sex pheromone CAD1.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OC NCBI_TaxID=151;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=85051889; PubMed=6437872; DOI=10.1016/0014-5793(84)81248-X;
RA Mori M., Sagakami Y., Narita M., Isegai A., Fujino M., Kitada C.,
RA Craig R.A., Clevellel D.B., Suzuki A.;
RT "Isolation and structure of the bacterial sex pheromone, CAD1, that
RT induces plasmid transfer in Streptococcus faecalis."
RL FEBS Lett. 178:97-100(1984).
CC -1- FUNCTION: Involved in the conjugative transfer of the hemolysin
CC plasmid PAD1.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 KM Direct protein sequencing; Phenone.
 SQ SEQUENCE 8 AA; 819 MW; 04DD732C735B9C7 CRC64;
 Query Match 5.2%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.2e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 FSL 35
 DB 2 FSL 4
 RESULT 12
 FARI_PANRE STANDARD; PRT; 8 AA.
 ID P41872;
 AC P41872;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE FMRamide-like neuropeptide PPI (SDPFLRF-amide).
 OS Panagrellus redivivus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
 OX NCBI_TaxId=6233;
 RN [1]
 RP PROTEIN SEQUENCE.
 RX MEDLINE=33027659; PubMed=1408999; DOI=10.1016/0196-9781(92)90098-N;
 RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,
 RA Garrison R.D., Williams J.F., Friedman A.R.;
 RT "Two FMRamide-like peptides from the free-living nematode Panagrellus
 RT redivivus.";
 RL Peptides 13:209-214(1992).
 CC -1- FUNCTION: Myoactive.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Nerve cords and paired groups of cells located
 CC caudally to the base of the pharynx.
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
 CC family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 KM Annotation; Direct protein sequencing; Neuropeptide.
 FT MOD RES 8 Phenylalanine amide.
 SQ SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;
 Query Match 5.2%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.2e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 29 FSL 31
 DB 5 FSL 7
 RESULT 13
 FARI_PENMO STANDARD; PRT; 8 AA.
 ID P83316;
 AC P83316;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE FMRamide-like neuropeptide FLPI (GPRNPLRF-amide).
 OS Penaeus monodon (Penaeid shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;

OC Penaeidae; Penaeus.
 OX NCBI_TaxId=6687;
 RN [1]
 RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Eyestalk;
 RX MEDLINE=21956277; PubMed=11959015; DOI=10.1016/S1096-4959(01)00499-7;
 RA Sithigorngul P., Pupum J., Krungkarn C., Longyant S.,
 RA Chaivuthangkura P., Sithigorngul W., Petsom A.;
 RT "Seven novel FMRamide-like neuropeptide sequences from the eyestalk
 RT of the giant tiger prawn Penaeus monodon.";
 RL Comp. Biochem. Physiol. 131B:325-337(2002).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=1024.8; METHOD=MALDI; RANGE=1-8; NOTE=Ref.1.
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
 CC family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR GO:0007218; P-neuropeptide signaling pathway; TMS.
 KW Annotation; Direct protein sequencing; Neuropeptide.
 FT MOD RES 8 Phenylalanine amide.
 SQ SEQUENCE 8 AA; 1024 MW; 72D40729C4540A8 CRC64;
 Query Match 5.2%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.2e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 29 FSL 31
 DB 5 FSL 7
 RESULT 14
 FARI_MACRS STANDARD; PRT; 8 AA.
 ID P83275;
 AC P83275;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE FMRamide-like neuropeptide FLPI (ADKNPLRF-amide).
 OS Macrobrachium rosenbergii (Giant fresh water prawn).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;
 OC Palaemonoidea; Palaemonidae; Macrobrachium.
 OX NCBI_TaxId=79674;
 RN [1]
 RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Eyestalk;
 RA Sithigorngul P., Sarathongkum W., Jaidechaoey S., Longyant S.,
 RA Sithigorngul W.;
 RT "Novel FMRamide-like neuropeptides from the eyestalk of the giant
 RT freshwater prawn Macrobrachium rosenbergii.";
 RL Comp. Biochem. Physiol. 120B:587-595(1998).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=1009.4; METHOD=MALDI; RANGE=1-8; NOTE=Ref.1.
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
 CC family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR GO:0007218; P-neuropeptide signaling pathway; TMS.
 KW Annotation; Direct protein sequencing; Neuropeptide.
 FT MOD RES 8 Phenylalanine amide.
 SQ SEQUENCE 8 AA; 1010 MW; 9CD40729C4433AD CRC64;

Query Match 5.2%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.2e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 FLR 31
 |||
 Db 5 FLR 7

RESULT 15

PAR3_HOMAM STANDARD; PRT; 8 AA.

AC P41486; 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE FMRFamide-like neuropeptide 3 (FLI 3) (P2).
 OS Homarus americanus (American lobster).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 OC Nephropoidea; Nephropidae; Homarus.
 OX NCBI_TaxID=6706;
 RN [1]

PROTEIN SEQUENCE.

RC TISSUE-Pericardial organs;
 RX MEDLINE=8816164; PubMed=3429714;
 RA Trimmer B.A., Kobliarski L.A., Kravitz E.A.;
 RT "Purification and characterization of FMRFamide-like immunoreactive
 RT substances from the lobster nervous system: isolation and sequence
 RT analysis of two closely related peptides.";
 RL J. Comp. Neurol. 265:16-26(1987).

CC -1 SUBCELLULAR LOCATION: Secreted.

CC -1 MISCELLANEOUS: Pericardial organs release this peptide with 100 mM
 potassium in the presence of calcium.

CC -1 SIMILARITY: Belongs to the PARP (FMRFamide related peptide)
 family.

CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC -----
 CC Amidation; Direct protein sequencing; Neuropeptide.

KW MOD RES 8 Phenylalanine amide.

FT SEQUENCE 8 AA; 1054 MW; C6D40729C4540AB5 CRC64;

Query Match 5.2%; Score 3; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.2e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 FLR 31
 |||
 Db 5 FLR 7

Search completed: December 13, 2005, 08:25:36
 Job time : 228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:37 ; Search time 46 Seconds

(without alignments)
104.243 Million cell updates/sec

Title: US-10-031-158b-14

Perfect score: 58

Sequence: 1 MGFPPSPPLFFQLKQSS.....RYGKRATRPWPRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 0

Total number of hits satisfying chosen parameters: 62763

Minimum DB seq length: 8
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RB.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	9	1	US-08-621-803-221
2	5	8.6	9	1	US-08-621-259A-213
3	5	8.6	9	2	US-08-159-339A-343
4	5	8.6	9	2	US-08-159-339A-405
5	5	8.6	9	2	US-08-217-352-221
6	5	8.6	9	2	US-09-677-664B-213
7	5	8.6	9	2	US-09-792-480-18
8	5	8.6	10	1	US-08-248-839C-135
9	5	8.6	10	2	US-09-648-400A-18
10	5	8.6	10	2	US-09-839-329-4
11	5	8.6	10	2	US-10-209-421-18
12	5	8.6	10	2	US-10-007-761-9
13	5	8.6	10	4	PCT-US96-01735-4
14	4	6.9	8	1	US-08-037-579A-6
15	4	6.9	8	1	US-08-244-855-6
16	4	6.9	8	1	US-08-529-190B-70
17	4	6.9	8	1	US-08-529-190B-71
18	4	6.9	8	1	US-08-922-267A-68
19	4	6.9	8	2	US-08-601-184-6
20	4	6.9	8	2	US-08-444-818-262
21	4	6.9	8	2	US-08-444-818-263
22	4	6.9	8	2	US-08-444-818-264
23	4	6.9	8	2	US-08-296-791-8
24	4	6.9	8	2	US-09-419-826-39
25	4	6.9	8	2	US-08-975-614-3
26	4	6.9	8	2	US-09-648-400A-10
27	4	6.9	8	2	US-08-475-955-193

28	4	6.9	8	2	US-08-475-955-194	Sequence 194, App
29	4	6.9	8	2	US-08-475-955-195	Sequence 195, App
30	4	6.9	8	2	US-08-475-955-196	Sequence 196, App
31	4	6.9	8	2	US-08-475-955-197	Sequence 197, App
32	4	6.9	8	2	US-08-475-955-204	Sequence 204, App
33	4	6.9	8	2	US-09-839-996-8	Sequence 8, Appl1
34	4	6.9	8	2	US-09-792-480-9	Sequence 9, Appl1
35	4	6.9	8	2	US-10-083-889-18	Sequence 18, Appl1
36	4	6.9	8	2	US-10-080-505-54	Sequence 54, Appl1
37	4	6.9	8	2	US-10-209-421-10	Sequence 10, Appl1
38	4	6.9	8	2	US-09-308-140-8	Sequence 8, Appl1
39	4	6.9	8	2	US-10-645-655-8	Sequence 8, Appl1
40	4	6.9	8	2	US-09-856-070-7	Sequence 7, Appl1
41	4	6.9	8	2	US-09-496-391-26	Sequence 26, Appl1
42	4	6.9	8	2	US-10-028-056-25	Sequence 25, Appl1
43	4	6.9	8	2	US-07-867-819D-139	Sequence 139, App
44	4	6.9	8	2	US-07-867-819D-141	Sequence 141, App
45	4	6.9	8	2	US-07-867-819D-142	Sequence 142, App

ALIGNMENTS

RESULT 1
US-08-621-803-221
Sequence 221, Application US/08621803
Patent No. 5851802
GENERAL INFORMATION:
APPLICANT: Belter, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
FUSION OF INVENTION: Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/621,803
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Borum, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO.: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Xmp.382:
NAME/KEY: Modified-site
LOCATION: C-Terminus
OTHER INFORMATION: /label= Amidation
OTHER INFORMATION: /note= "The C-Terminus is Amidated."
US-08-621-803-221
Query Match 8.6%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQULK 17
|||||
Db 4 LQULK 8

RESULT 2
US-08-621-259A-213
Sequence 213, Application US/08621259A
Patent No. 5858974
GENERAL INFORMATION:
APPLICANT: Little II, Roger G
APPLICANT: Lim, Edward
APPLICANT: Padem, Mitchell B.
TITLE OF INVENTION: Anti-Fungal Peptides
NUMBER OF SEQUENCES: 252
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrew, Held & Malloy, Ltd.
STREET: 500 West Madison Street
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/621,259A
FILING DATE: 21-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/504,841
FILING DATE: 20-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11021US02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX:
INFORMATION FOR SEQ ID NO: 213:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: "XMP.382:
NAME/KEY: Modified-site
LOCATION: C-terminus
OTHER INFORMATION: //label= Amidation
OTHER INFORMATION: /note= "The C-terminus is Amidated."
US-08-621-259A-213

Query Match 8.6%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQULK 17
|||||
Db 4 LQULK 8

RESULT 3
US-08-159-339A-343
Sequence 343, Application US/08159339A
Patent No. 6037135
GENERAL INFORMATION:

APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Sette, Alessandro
APPLICANT: Celis, Esben
TITLE OF INVENTION: HLA Binding peptides and Their
TITLE OF INVENTION: Uses
NUMBER OF SEQUENCES: 1254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,339A
FILING DATE: 29-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauer
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-005030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 343:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-159-339A-343

Query Match 8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 SLMLL 38
|||||
Db 1 SLMLL 5

RESULT 4
US-08-159-339A-405
Sequence 405, Application US/08159339A
Patent No. 6037135
GENERAL INFORMATION:
APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Sette, Alessandro
APPLICANT: Celis, Esben
TITLE OF INVENTION: HLA Binding peptides and Their
TITLE OF INVENTION: Uses
NUMBER OF SEQUENCES: 1254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,339A
FILING DATE: 29-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-005030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 405:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-159-339A-405

Query Match 8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SLML 38
Db 3 SLML 7

RESULT 5
US-09-217-352-221
Sequence 221, Application US/09217352
Patent No. 6274344
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6400 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/217,352
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: "XMP.382:
FEATURE:
NAME/KEY: Modified-site
LOCATION: C-Terminus
OTHER INFORMATION: /label= Amidation
OTHER INFORMATION: /note= "The C-Terminus is Amidated."
US-09-217-352-221

Query Match 8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17
Db 4 LQLLK 8

RESULT 6
US-09-677-664B-213
Sequence 213, Application US/09677664B
Patent No. 6664231
GENERAL INFORMATION:
APPLICANT: Little II, Roger G
Lim, Edward
Padem, Mitchell B.
TITLE OF INVENTION: Anti-Fungal Peptides
NUMBER OF SEQUENCES: 257
CORRESPONDENCE ADDRESSES:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/677,664B
FILING DATE: 07-Mar-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/227,659
FILING DATE: 08-Jan-1999
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11021US06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/775-8000
TELEFAX: 312/775-8100
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 213:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

```

;
; FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: "XMP.382:"
;
; FEATURE:
;   NAME/KEY: Modified-site
;   LOCATION: C-Terminus
;   OTHER INFORMATION: /label= Amidation
;   /note= "The C-Terminus is Amidated."
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 213:
US-09-677-664B-213

Query Match      8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      13 LQLLK 17
Db      4 LQLLK 8

RESULT 7
US-09-792-480-18
; Sequence 18, Application US/09792480
; Patent No. 6669951
; GENERAL INFORMATION:
;   APPLICANT: Rothbard, Jonathan B.
;   APPLICANT: Wender, Paul A.
;   APPLICANT: McGrane, P. Leo
;   APPLICANT: Siesta, Lalitha V.S.
;   APPLICANT: Kirschberg, Thorsten A.
;   APPLICANT: Cellgate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000230US
; CURRENT APPLICATION NUMBER: US/09/792,480
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 09/648,400
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region tat-49-57
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa = fluorescein linked to amino group of
;   OTHER INFORMATION: aminohexanoic acid (Fl-ahx) attached to the
;   OTHER INFORMATION: N-terminal amino group of Arg
US-09-792-480-18

Query Match      8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      43 KKRRA 47
Db      2 KKRRA 6

RESULT 8
US-08-248-839C-135
; Sequence 135, Application US/08248839C
; Patent No. 5843702
; GENERAL INFORMATION:
;   APPLICANT: McComell, David
;   APPLICANT: Devine, Kevin
```

```

;
; APPLICANT: O'Kane, Charles
; TITLE OF INVENTION: A Gene Expression System
; NUMBER OF SEQUENCES: 185
; CORRESPONDENCE ADDRESSES:
;   ADDRESSEE: No. 58437020 No. 5843702disk of No. 5843702th America, Inc.
;   STREET: 405 Lexington Avenue
;   CITY: New York
;   STATE: New York
;   COUNTRY: USA
;   ZIP: 10174-6401
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/248,839C
;   FILING DATE: 25-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Gregg, Valeta A.
;   REGISTRATION NUMBER: 35,127
;   REFERENCE/DOCKET NUMBER: 3614.214-US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 212-867-0123
;   TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 10 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-248-839C-135

Query Match      8.6%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      34 SLMLL 38
Db      3 SLMLL 7

RESULT 9
US-09-648-400A-18
; Sequence 18, Application US/09648400A
; Patent No. 6593292
; GENERAL INFORMATION:
;   APPLICANT: Rothbard, Jonathan B.
;   APPLICANT: Wender, Paul A.
;   APPLICANT: McGrane, P. Leo
;   APPLICANT: Siesta, Lalitha V.S.
;   APPLICANT: Kirschberg, Thorsten A.
;   APPLICANT: Cellgate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000210US
; CURRENT APPLICATION NUMBER: US/09/648,400A
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region tat-49-57
; FEATURE:
```

NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa = fluorescein linked to amino group of
OTHER INFORMATION: aminohexanoic acid (Pl-ahx)
US-09-648-400A-18

Query Match
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 GKRR 47
Db 3 KKRR 7

RESULT 10
US-09-839-329-4
Sequence 4, Application US/09839329
Patent No. 6664040
GENERAL INFORMATION:
APPLICANT: Michael P. Sherman
APPLICANT: Warner C. Greene
APPLICANT: Carlos M.C. de No. 6664040hna
APPLICANT: Ulrich Schubert
APPLICANT: Peter Henklein
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
FILE REFERENCE: G&C 30448.91-US-U2
CURRENT APPLICATION NUMBER: US/09/839,329
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: 60/206,610
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/267,827
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Human immunodeficiency virus
US-09-839-329-4

Query Match
Best Local Similarity 8.6%; Score 5; DB 2; Length 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 GKRR 46
Db 2 GKRR 6

RESULT 11
US-10-209-421-18
Sequence 18, Application US/10209421
Patent No. 6759387
GENERAL INFORMATION:
APPLICANT: Rothbard, Jonathan B.
APPLICANT: Wender, Paul A.
APPLICANT: McGrane, P. Leo
APPLICANT: Sista, Lalitha V.S.
APPLICANT: Kirschberg, Thorsten A.
APPLICANT: Celigateg, Inc.
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
FILE REFERENCE: 019801-000211US
CURRENT APPLICATION NUMBER: US/10/209,421
PRIOR FILING DATE: 2002-07-30
PRIOR APPLICATION NUMBER: US 60/150,510
PRIOR FILING DATE: 1999-08-24
PRIOR APPLICATION NUMBER: US 09/648,400
PRIOR FILING DATE: 2000-08-24

NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 18
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:A-54
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
OTHER INFORMATION: region Tat-49-57

NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa = fluorescein linked to amino group of
OTHER INFORMATION: aminohexanoic acid (Pl-ahx)
US-10-209-421-18

Query Match
Best Local Similarity 8.6%; Score 5; DB 2; Length 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KKRR 47
Db 3 KKRR 7

RESULT 12
US-10-007-761-9
Sequence 9, Application US/10007761
Patent No. 6855693
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: Peptides for Activation and Inhibition
FILE REFERENCE: 58600-8208-US00
CURRENT APPLICATION NUMBER: US/10/007,761
CURRENT FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: US 60/262,060
PRIOR FILING DATE: 2001-01-18
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Tat-derived carrier peptide
US-10-007-761-9

Query Match
Best Local Similarity 8.6%; Score 5; DB 2; Length 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 GKRR 46
Db 2 GKRR 6

RESULT 13
PCT-US96-01735-4
Sequence 4, Application PC/TUS9601735
GENERAL INFORMATION:
APPLICANT: Marks, Andrew R.
TITLE OF INVENTION: HUMAN T CELL INOSITOL 1,4,5,-TRISPHOSPHATE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brunbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228


```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01735
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/386,039
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, Lisa B
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: A30042 - 165/30555
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2628
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; FEATURE:
; PCT-US96-01735-4

Query Match      8.6%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred.No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      12 FLOLL 16
      |||||
Db      4 FLOLL 8

RESULT 14
US-08-037-579A-6
; Sequence 6, Application US/08037579A
; Patent No. 5552537
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Max, Edward E
; APPLICANT: Saxon, Andrew
; TITLE OF INVENTION: IGE ISOFORMS AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOERBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/037,579A
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-57950/BIR UCLA-233
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277229 FHT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-037-579A-6

Query Match      6.9%; Score 4; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 4.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 SSRR 22
      |||||
Db      2 SSRR 5

RESULT 15
US-08-244-855-6
; Sequence 6, Application US/08244855
; Patent No. 5762906
; GENERAL INFORMATION:
; APPLICANT: Creighton, Andrew M.
; TITLE OF INVENTION: FURTHER IMPROVEMENTS RELATING TO
; TITLE OF INVENTION: RADIO LABELLING OF PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHIE
; STREET: 8TH FLOOR, 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,855
; FILING DATE: 16-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 604-285
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4005
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-244-855-6

Qy      45 RRAT 48
      |||||
Db      2 RRAT 5

Search completed: December 13, 2005, 08:26:27
```

Wed Dec 14 08:46:43 2005

us-10-031-158b-14.oliszm.rat

Page 7

Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:25:42 ; Search time 159 Seconds
(Without alignments)
152.416 Million cell updates/sec

Title: US-10-031-158b-14

Perfect score: 58
Sequence: 1 MCMFPSPPLFFFLQLKQSS.....RYGKKRRATRPWDRRGRTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size : 0

Total number of hits satisfying chosen parameters: 151674

Minimum DB seq length: 8
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	8	US-10-137-867-292	Sequence 292, App
2	5	8.6	9	US-09-765-527-221	Sequence 221, App
3	5	8.6	9	US-09-792-480-18	Sequence 18, Appl
4	5	8.6	9	US-09-870-216C-3	Sequence 3, Appli
5	5	8.6	9	US-10-017-327-3	Sequence 3, Appli
6	5	8.6	9	US-10-144-188-2	Sequence 2, Appli
7	5	8.6	10	US-09-839-329-4	Sequence 4, Appli
8	5	8.6	10	US-09-572-404B-238	Sequence 238, App
9	5	8.6	10	US-09-572-404B-1444	Sequence 1444, Ap
10	5	8.6	10	US-09-572-404B-3212	Sequence 3212, Ap
11	5	8.6	10	US-10-007-761-9	Sequence 9, Appli
12	5	8.6	10	US-10-083-960-18	Sequence 18, Appl
13	5	8.6	10	US-10-209-421-18	Sequence 18, Appl
14	5	8.6	10	US-10-168-789A-26	Sequence 26, Appl
15	5	8.6	10	US-10-168-789A-27	Sequence 27, Appl
16	5	8.6	10	US-10-421-548-7	Sequence 7, Appli
17	5	8.6	10	US-10-421-503-65	Sequence 65, Appli
18	5	8.6	10	US-10-843-731-9	Sequence 9, Appli
19	5	8.6	10	US-10-936-237-47	Sequence 47, Appl
20	5	8.6	10	US-10-513-003-65	Sequence 65, Appl
21	4	6.9	8	US-09-815-837-114	Sequence 114, App
22	4	6.9	8	US-09-804-682-170	Sequence 170, App
23	4	6.9	8	US-09-792-480-9	Sequence 9, Appli
24	4	6.9	8	US-09-756-875-10	Sequence 10, Appl
25	4	6.9	8	US-09-756-875-11	Sequence 11, Appl
26	4	6.9	8	US-09-756-875-20	Sequence 20, Appl
27	4	6.9	8	US-09-756-875-21	Sequence 21, Appl

28	4	6.9	8	US-09-756-875-22	Sequence 22, Appl
29	4	6.9	8	US-09-756-875-23	Sequence 23, Appl
30	4	6.9	8	US-09-756-875-24	Sequence 24, Appl
31	4	6.9	8	US-09-756-875-29	Sequence 29, Appl
32	4	6.9	8	US-09-017-743C-7	Sequence 7, Appli
33	4	6.9	8	US-09-943-123-19	Sequence 19, Appl
34	4	6.9	8	US-09-839-996-8	Sequence 8, Appli
35	4	6.9	8	US-09-726-470A-76	Sequence 76, Appl
36	4	6.9	8	US-09-726-470A-137	Sequence 137, App
37	4	6.9	8	US-09-726-470A-244	Sequence 244, App
38	4	6.9	8	US-09-792-861A-15	Sequence 15, Appl
39	4	6.9	8	US-09-770-875-15	Sequence 15, Appl
40	4	6.9	8	US-09-876-904A-353	Sequence 353, App
41	4	6.9	8	US-09-876-904A-408	Sequence 408, App
42	4	6.9	8	US-09-876-904A-417	Sequence 417, App
43	4	6.9	8	US-09-876-904A-419	Sequence 419, App
44	4	6.9	8	US-09-796-076-10	Sequence 10, Appl
45	4	6.9	8	US-10-028-056-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-10-137-867-292
; Sequence 292, Application US/10137867
; Publication No. US20030207349A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Tamas, Daniel
; APPLICANT: Tamas, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C146
; CURRENT APPLICATION NUMBER: US/10137,867
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 292
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-867-292

Query Match 8.6%; Score 5; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 9 LFFFL 13
Db 1 LFFFL 5

RESULT 2
US-09-765-527-221
; Sequence 221, Application US/09765527
; Patent No. US20020006638A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.

TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerestein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: m1sc_feature
OTHER INFORMATION: "XMP.382:
FEATURE:
NAME/KEY: Modified-site
LOCATION: C-Terminus
OTHER INFORMATION: //label= Amidation
/note= "The C-Terminus is Amidated."
SEQUENCE DESCRIPTION: SEQ ID NO: 221:
US-09-765-527-221
Query Match 8.6%; Score 5; DB 3; Length 9;
Best Local Similarity 100.0%; Pred.No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 LQLLK 17
Db 4 LQLLK 8
RESULT 3
US-09-792-480-18
Sequence 18, Application US/09792480
Patent No. US20020127198A1
GENERAL INFORMATION:
APPLICANT: Rothbard, Jonathan B.
APPLICANT: Wender, Paul A.
APPLICANT: McGrane, P. Leo
APPLICANT: Sista, Lalitha V.S.
APPLICANT: Kirschberg, Thorsten A.
APPLICANT: Cellgate, Inc.
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
TITLE OF INVENTION: Across and Into Epithelial Tissues
FILE REFERENCE: 019801-000230US
CURRENT APPLICATION NUMBER: US/09/792,480
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: US 09/648,400
PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/150,510
PRIOR FILING DATE: 1999-08-24
NUMBER OF SEQ ID NOS: 57
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 18
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:A-54
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
OTHER INFORMATION: region Tat-49-57
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa = fluorescein linked to amino group of
OTHER INFORMATION: aminohexanoic acid (PI-ahx) attached to the
OTHER INFORMATION: N-terminal amino group of Arg
US-09-792-480-18
Query Match 8.6%; Score 5; DB 3; Length 9;
Best Local Similarity 100.0%; Pred.No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 43 KKRA 47
Db 2 KKRA 6
RESULT 4
US-09-870-216C-3
Sequence 3, Application US/09870216C
Publication No. US20040138135A1
GENERAL INFORMATION:
APPLICANT: Charles A. Nicolette
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER
FILE REFERENCE: 68126881210100
CURRENT APPLICATION NUMBER: US/09/870,216C
CURRENT FILING DATE: 2001-05-30
PRIOR FILING DATE: 2000-05-31
PRIOR APPLICATION NUMBER: 60/209,391
PRIOR FILING DATE: 2000-08-17/226,256
PRIOR APPLICATION NUMBER: 60/257,008
PRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-09-870-216C-3
Query Match 8.6%; Score 5; DB 3; Length 9;
Best Local Similarity 100.0%; Pred.No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 FLQLL 16
Db 1 FLQLL 5
RESULT 5
US-10-017-327-3
Sequence 3, Application US/10017327
Publication No. US20020155472A1
GENERAL INFORMATION:
APPLICANT: Charles A. Nicolette
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
TITLE OF INVENTION: METHODS FOR USING SAME
FILE REFERENCE: GZ 2101.20
CURRENT APPLICATION NUMBER: US/10/017,327
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 11

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-327-3
```

```
Query Match      8.6%; Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12 FLOQL 16
        |||||
Db       1 FLOQL 5
```

```
RESULT 6
US-10-144-188-2
; Sequence 2, Application US/10144188
; Publication No. US20030170212A1
; GENERAL INFORMATION:
; APPLICANT: Cal, Zeiling
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; APPLICANT: Shi, Weixing
; APPLICANT: Kong, Yan
; APPLICANT: Degraw, Juli
; TITLE OF INVENTION: Ex-Vivo Priming For Generating Cytotoxic T Lymphocytes Specific
; TITLE OF INVENTION: For No. US20030170212A1-Tumor Antigens To Treat Autoimmune And A
; FILE REFERENCE: PRI0010 ORT-1627
; CURRENT APPLICATION NUMBER: US/10/144,188
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/291,300
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide antigen
US-10-144-188-2
```

```
Query Match      8.6%; Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 PPSPL 9
        |||||
Db       2 PPSPL 6
```

```
RESULT 7
US-09-839-329-4
; Sequence 4, Application US/09839329
; Publication No. US2002002027A1
; GENERAL INFORMATION:
; APPLICANT: Michael P. Sherman
; APPLICANT: Warner C. Greene
; APPLICANT: Carlos M.C. de No. 6664040hna
; APPLICANT: Ulrich Schubert
; APPLICANT: Peter Henklein
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
; TITLE OF INVENTION: A MOLECULE INTO A CELL
; FILE REFERENCE: G&C 30448-91-US-U2
; CURRENT APPLICATION NUMBER: US/09/839,329
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/206,610
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/267,827
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 5
```

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human immunodeficiency virus
US-09-839-329-4
```

```
Query Match      8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      42 GKRR 46
        |||||
Db       2 GKRR 6
```

```
RESULT 8
US-09-572-404B-238
; Sequence 238, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 238
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in MTNR1B at 351-360 and may interact with Sequer
US-09-572-404B-238
```

```
Query Match      8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      20 SRRL 24
        |||||
Db       1 SRRL 5
```

```
RESULT 9
US-09-572-404B-1444
; Sequence 1444, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 1444
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in NTAK at 681-690 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-1444
```

```
Query Match      8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 53 PRGRT 57
Db 6 PRGRT 10

RESULT 10
US-09-572-404B-3212
; Sequence 3212, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 3212
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Sequence located in HOXA9 OR HOXA10 at 369-378 and may interact w/1
US-09-572-404B-3212

Query Match 8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 RRATR 49
Db 4 RRATR 8

RESULT 11
US-10-007-761-9
; Sequence 9, Application US/10007761
; Publication No. US20020150984A1
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: Peptides for Activation and Inhibition
; FILE REFERENCE: Of delta-PKC
; FILE REFERENCE: 58600-8208.US00
; CURRENT APPLICATION NUMBER: US/10/007,761
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/262,060
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Tat-derived carrier peptide
US-10-007-761-9

Query Match 8.6%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 GKRR 46
Db 2 GKRR 6

RESULT 12
US-10-083-960-18
; Sequence 18, Application US/10083960
; Publication No. US20030022831A1
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.

APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing
; TITLE OF INVENTION: Drug Delivery Across and Into Ocular Tissues
; FILE REFERENCE: 019801-000240US
; CURRENT APPLICATION NUMBER: US/10/083,960
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: US 09/648,400
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 09/792,480
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: A-54 alanine-substituted analog of Tat-49-57
; NAME/KEY: MOD_RES
; LOCATION: (1)...(1)
; OTHER INFORMATION: Xaa = fluorescein conjugated aminohexanoic acid
; OTHER INFORMATION: (Fl-ahx)
US-10-083-960-18

Query Match 8.6%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 KKRA 47
Db 3 KKRA 7

RESULT 13
US-10-209-421-18
; Sequence 18, Application US/10209421
; Publication No. US20030083256A1
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000211US
; CURRENT APPLICATION NUMBER: US/10/209,421
; CURRENT FILING DATE: 2002-07-30
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: US 09/648,400
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:A-54
; OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
; OTHER INFORMATION: region Tat-49-57
; NAME/KEY: MOD_RES
; LOCATION: (1)

; OTHER INFORMATION: Xaa = fluorescein linked to amino group of
; OTHER INFORMATION: aminohexanoic acid (Fl-ahx)
US-10-209-421-18

Query Match 8.6%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRRA 47
DB 3 KKRRA 7

RESULT 14
US-10-168-789A-26.

; Sequence 26, Application US/10168789A
; Publication No. US20030148943A1
; GENERAL INFORMATION:
; APPLICANT: ITOH, Yasuaki
; APPLICANT: NISHI, Kazumori
; APPLICANT: KITADA, Chieko
; APPLICANT: INATOMI, No. US20030148943A1uhiro
; TITLE OF INVENTION: No. US20030148943A1el Tachykinin-like Polypeptides and Use Thereof
; FILE REFERENCE: 2680USOP
; CURRENT APPLICATION NUMBER: US/10/168, 789A
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: PCT/JP00/09083
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: JP 11-362638
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: JP 12-066714
; PRIOR FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 64
; SEQ ID NO 26
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide
US-10-168-789A-26

Query Match 8.6%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LEHTF 27
DB 2 LEHTF 6

RESULT 15
US-10-168-789A-27.

; Sequence 27, Application US/10168789A
; Publication No. US20030148943A1
; GENERAL INFORMATION:
; APPLICANT: ITOH, Yasuaki
; APPLICANT: NISHI, Kazumori
; APPLICANT: KITADA, Chieko
; APPLICANT: INATOMI, No. US20030148943A1uhiro
; TITLE OF INVENTION: No. US20030148943A1el Tachykinin-like Polypeptides and Use Thereof
; FILE REFERENCE: 2680USOP
; CURRENT APPLICATION NUMBER: US/10/168, 789A
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: PCT/JP00/09083
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: JP 11-362638
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: JP 12-066714
; PRIOR FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 64
; SEQ ID NO 27
; LENGTH: 10
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Polypeptide
; NAME/KEY: PEPTIDE
; LOCATION: (01)..(01)
; OTHER INFORMATION: Xaa means pyroglutamic acid
US-10-168-789A-27

Query Match 8.6%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LEHTF 27
DB 2 LEHTF 6

Search completed: December 13, 2005, 08:39:30
Job time : 160 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:26:32 / Search time 11 Seconds

(without alignments)
29,445 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58
Sequence: 1 MCMFPPSPLEFFILQKQSS.....RYIGKKRRATRFMDPRGRTP 58

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 32527 seqs, 5584426 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3786

Minimum DB seq length: 8
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database:

Published Applications AA New:
1: /cgn2_6/ptcodata/2/pubppa/US09_NEW_PUB pep:*
2: /cgn2_6/ptcodata/2/pubppa/US06_NEW_PUB pep:*
3: /cgn2_6/ptcodata/2/pubppa/US07_NEW_PUB pep:*
4: /cgn2_6/ptcodata/2/pubppa/US08_NEW_PUB pep:*
5: /cgn2_6/ptcodata/2/pubppa/PC7_NEW_PUB pep:*
6: /cgn2_6/ptcodata/2/pubppa/US10_NEW_PUB pep:*
7: /cgn2_6/ptcodata/2/pubppa/US11_NEW_PUB pep:*
8: /cgn2_6/ptcodata/2/pubppa/US60_NEW_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	10	US-11-119-098-8	Sequence 8, App1
2	4	6.9	8	US-10-914-842A-26	Sequence 26, App1
3	4	6.9	9	US-10-499-715-15	Sequence 15, App1
4	4	6.9	9	US-10-499-715-15	Sequence 16, App1
5	4	6.9	9	US-11-176-868-1	Sequence 1, App1
6	4	6.9	9	US-11-097-864-144	Sequence 144, App
7	4	6.9	9	US-11-097-912-144	Sequence 144, App
8	4	6.9	10	US-10-444-562-1	Sequence 1, App1
9	4	6.9	10	US-11-152-747-28	Sequence 28, App1
10	4	6.9	10	US-11-032-794-53	Sequence 53, App1
11	4	6.9	10	US-11-097-864-446	Sequence 446, App
12	4	6.9	10	US-11-097-864-446	Sequence 511, App
13	4	6.9	10	US-11-097-912-446	Sequence 446, App
14	4	6.9	10	US-11-097-912-511	Sequence 511, App
15	4	6.9	10	US-11-102-432-1	Sequence 1, App1
16	3	5.2	8	US-10-997-066-22	Sequence 22, App1
17	3	5.2	8	US-10-839-966-6	Sequence 6, App1
18	3	5.2	8	US-10-467-657-1372	Sequence 7372, Ap
19	3	5.2	8	US-10-467-657-8729	Sequence 8729, Ap
20	3	5.2	8	US-10-467-657-8769	Sequence 8769, Ap
21	3	5.2	8	US-10-467-657-8797	Sequence 8797, Ap
22	3	5.2	8	US-10-914-842A-17	Sequence 17, App1
23	3	5.2	8	US-10-416-047-5	Sequence 5, App1
24	3	5.2	8	US-11-082-251-14	Sequence 14, App1
25	3	5.2	8	US-11-096-706-208	Sequence 208, App

26	3	5.2	8	US-11-176-868-17	Sequence 17, App1
27	3	5.2	8	US-11-032-498-24	Sequence 24, App1
28	3	5.2	8	US-11-152-747-32	Sequence 32, App1
29	3	5.2	8	US-11-054-515-3141	Sequence 3141, Ap
30	3	5.2	8	US-11-054-515-3214	Sequence 3214, Ap
31	3	5.2	8	US-11-058-727-98	Sequence 98, App1
32	3	5.2	8	US-11-058-735-8	Sequence 8, App1
33	3	5.2	8	US-11-058-735-9	Sequence 9, App1
34	3	5.2	8	US-11-058-735-10	Sequence 10, App1
35	3	5.2	8	US-11-108-389-98	Sequence 98, App1
36	3	5.2	8	US-11-101-287-152	Sequence 152, App
37	3	5.2	8	US-11-115-922-213	Sequence 213, App
38	3	5.2	9	US-10-997-066-8	Sequence 8, App1
39	3	5.2	9	US-10-997-066-10	Sequence 10, App1
40	3	5.2	9	US-10-499-715-17	Sequence 17, App1
41	3	5.2	9	US-10-499-715-18	Sequence 18, App1
42	3	5.2	9	US-10-981-873-26	Sequence 26, App1
43	3	5.2	9	US-10-073-301A-10	Sequence 10, App1
44	3	5.2	9	US-10-952-535A-44	Sequence 44, App1
45	3	5.2	9	US-10-467-657-9002	Sequence 9002, Ap

ALIGNMENTS

```
RESULT 1
US-11-119-098-8
; Sequence 8, Application US/1119098
; Publication No. US20050267030A1
; GENERAL INFORMATION:
; APPLICANT: Taao, Philip S.
; TITLE OF INVENTION: Use of deltaPKC Peptides for Modulation of Reactive Oxygen Species
; FILE REFERENCE: 58600-8213-US00
; CURRENT APPLICATION NUMBER: US/11/119, 098
; PRIOR FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: US 60/567,315
; PRIOR FILING DATE: 2004-04-30
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Tat-derived carrier peptide
US-11-119-098-8

Query Match      8.6% Score 5; DB 7; Length 10;
Best Local Similarity 100.0%; Pred.No. 7.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      42 GKRR 46
Db      2 GKRR 6

RESULT 2
US-10-914-842A-26
; Sequence 26, Application US/10914842A
; Publication No. US20050260730A1
; GENERAL INFORMATION:
; APPLICANT: FISCHER, PETER MARTIN
; TITLE OF INVENTION: CDK2/CYCLIN A CRYSTALS AND USES THEREOF
; FILE REFERENCE: CCI-032
; CURRENT APPLICATION NUMBER: US/10/914, 842A
; PRIOR FILING DATE: 2004-08-10
; PRIOR APPLICATION NUMBER: UK 0324465.4
; PRIOR FILING DATE: 2003-10-20
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 26
; LENGTH: 8
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-914-842A-26
```

```
Query Match
Best Local Similarity 100.0%; Score 4; DB 6; Length 8;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 44 KKRR 47
DB 3 KKRA 6
```

```
RESULT 3
US-10-499-715-15
; Sequence 15, Application US/10499715
; Publication No. US20050250717A1
; GENERAL INFORMATION:
; APPLICANT: BENAVAHU, Dafna
; APPLICANT: SHUR, Iritna
; TITLE OF INVENTION: CHROMATIN REMODELING PROTEIN AS A MARKER EXPRESSED BY STROMAL PRO
; FILE REFERENCE: BENAVAHU-1.1 PCT
; CURRENT APPLICATION NUMBER: US/10/499,715
; CURRENT FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-499-715-15
```

```
Query Match
Best Local Similarity 100.0%; Score 4; DB 6; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 43 KKRR 46
DB 1 KKRR 4
```

```
RESULT 4
US-10-499-715-16
; Sequence 16, Application US/10499715
; Publication No. US20050250717A1
; GENERAL INFORMATION:
; APPLICANT: BENAVAHU, Dafna
; APPLICANT: SHUR, Iritna
; TITLE OF INVENTION: CHROMATIN REMODELING PROTEIN AS A MARKER EXPRESSED BY STROMAL PRO
; FILE REFERENCE: BENAVAHU-1.1 PCT
; CURRENT APPLICATION NUMBER: US/10/499,715
; CURRENT FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-499-715-16
```

```
Query Match
Best Local Similarity 100.0%; Score 4; DB 6; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 43 KKRR 46
DB 1 KKRR 4
```

```
RESULT 5
US-11-176-868-1
; Sequence 1, Application US/1176868
; Publication No. US20050245454A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing Multiplication of HIV-1
; FILE REFERENCE: GGPJUSA
; CURRENT APPLICATION NUMBER: US/11/176,868
; CURRENT FILING DATE: 2005-07-08
; PRIOR APPLICATION NUMBER: US/10/323,013
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/10/114,176
; PRIOR FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-176-868-1
```

```
Query Match
Best Local Similarity 100.0%; Score 4; DB 7; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 43 KKRR 46
DB 2 KKRR 5
```

```
RESULT 6
US-11-097-864-144
; Sequence 144, Application US/11097864
; Publication No. US20050265924A1
; GENERAL INFORMATION:
; APPLICANT: Chailita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 161P2F10B
; FILE REFERENCE: 511582006205
; CURRENT APPLICATION NUMBER: US/11/097,864
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-097-864-144
```

```
Query Match
Best Local Similarity 100.0%; Score 4; DB 7; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 32 NFSL 35
DB 6 NFSL 9
```

```
RESULT 7
US-11-097-912-144
; Sequence 144, Application US/11097912
; Publication No. US20050265921A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 161P2F10B
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 511582006204
; CURRENT APPLICATION NUMBER: US/11/097,912
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-097-912-144

Query Match          6.9%; Score 4; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFSL 35
   ||||
Db 6 NFSL 9

RESULT 8
US-10-444-662-1
; Sequence 1, Application US/10444662
; Publication No. US20050250683A9
; GENERAL INFORMATION:
; APPLICANT: Mirus Corporation
; APPLICANT: Rozema, David
; APPLICANT: Wolff, Jon
; APPLICANT: Wakefield, Darren
; APPLICANT: Ekens, Kirk
; APPLICANT: Hagerstrom, James
; TITLE OF INVENTION: Reversible Modification of Membrane Interaction
; FILE REFERENCE: Mirus.035.01
; CURRENT APPLICATION NUMBER: US/10/444,662
; CURRENT FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-444-662-1

Query Match          6.9%; Score 4; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 46
   ||||
Db 3 KKRR 6

RESULT 9
US-11-152-747-28
; Sequence 28, Application US/11152747
; Publication No. US20050251881A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
```

```
; TITLE OF INVENTION: CAROTENOID KETOYLASE GENE
; FILE REFERENCE: CL-1849 US NA
; CURRENT APPLICATION NUMBER: US/11/152,747
; CURRENT FILING DATE: 2005-06-14
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis AN12
US-11-152-747-28

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LRYI 41
   ||||
Db 3 LRYI 6

RESULT 10
US-11-032-794-53
; Sequence 53, Application US/11032794
; Publication No. US20050265970A1
; GENERAL INFORMATION:
; APPLICANT: FRANZOSO, GUIDO
; APPLICANT: PAPA, SALVATORE
; APPLICANT: BUBICI, CONCETTA
; APPLICANT: DESMAELE, ENRICO
; APPLICANT: ZAZZERONI, FRANCESCA
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL FACTORS THAT BLOCK PROGRAMMED
; TITLE OF INVENTION: CELL DEATH OR APOPTOSIS BY TARGETING JNK
; FILE REFERENCE: 21459-97816
; CURRENT APPLICATION NUMBER: US/11/032,794
; CURRENT FILING DATE: 2005-01-10
; PRIOR APPLICATION NUMBER: 11/000,365
; PRIOR FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 60/526,231
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: 10/626,905
; PRIOR FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 10/263,330
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/328,811
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/326,492
; PRIOR FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 53
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-032-794-53

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 46
   ||||
Db 3 KKRR 6

RESULT 11
US-11-097-864-446
; Sequence 446, Application US/11097864
; Publication No. US20050265924A1
; GENERAL INFORMATION:
; APPLICANT: Challita-Bid, Pia M.
```

```

: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 161P2F10B
: TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006205
: CURRENT APPLICATION NUMBER: US/11/097,864
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 446
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-864-446
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 12
US-11-097-864-511
: Sequence 511, Application US/11097864
: Publication No. US20050265924A1
: GENERAL INFORMATION:
: APPLICANT: Chailita-Bid, Pia M.
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 161P2F10B
: TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006205
: CURRENT APPLICATION NUMBER: US/11/097,864
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 511
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-864-511
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 13
US-11-097-912-446
: Sequence 446, Application US/11097912
: Publication No. US20050265921A1
: GENERAL INFORMATION:
```

```

: APPLICANT: Chailita-Bid, Pia M.
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 161P2F10B
: TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006204
: CURRENT APPLICATION NUMBER: US/11/097,912
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 446
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-912-446
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 14
US-11-097-912-511
: Sequence 511, Application US/11097912
: Publication No. US20050265921A1
: GENERAL INFORMATION:
: APPLICANT: Chailita-Bid, Pia M.
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 161P2F10B
: TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006204
: CURRENT APPLICATION NUMBER: US/11/097,912
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 511
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-912-511
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 15
US-11-102-432-1
: Sequence 1, Application US/11102432
: Publication No. US20050265957A1
```

; GENERAL INFORMATION:
; APPLICANT: Mirus Bio
; APPLICANT: Monahan, Sean
; APPLICANT: Nader, Lisa
; TITLE OF INVENTION: Polymerized Formamides for Use in Delivery of Compounds to Cells
; FILE REFERENCE: Mirus.044.01
; CURRENT APPLICATION NUMBER: US/11/102,432
; CURRENT FILING DATE: 2005-04-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-11-102-432-1

Query Match 6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46
|||
Db 3 KKR 6

Search completed: December 13, 2005, 08:39:48
Job time : 12 secs

THIS PAGE BLANK (USPTO)